Wiley Finance Series

FOURTH EDITION

FIXED INCOME SECURITIES

Tools for Today's Markets

BRUCE TUCKMAN
ANGEL SERRAT

WILEY

Fixed Income Securities

Founded in 1807, John Wiley & Sons is the oldest independent publishing company in the United States. With offices in North America, Europe, Australia and Asia, Wiley is globally committed to developing and marketing print and electronic products and services for our customers' professional and personal knowledge and understanding.

The Wiley Finance series contains books written specifically for finance and investment professionals as well as sophisticated individual investors and their financial advisors. Book topics range from portfolio management to e-commerce, risk management, financial engineering, valuation, and financial instrument analysis, as well as much more.

For a list of available titles, visit our website at www.WileyFinance.com.

Fixed Income Securities

Tools for Today's Markets

Fourth Edition

BRUCE TUCKMAN
ANGEL SERRAT

WILEY

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data is Available:

ISBN 9781119835554 (Hardback)
ISBN 9781119835608 (ePDF)
ISBN 9781119835592 (ePub)
ISBN 9781119835622 (oBook)

Cover Design: Wiley
Cover Image: © MF3d/Getty Images

Contents

Preface

The goal of this book is to convey the institutional, conceptual, and quantitative frameworks used by sophisticated fixed income market practitioners. The overview chapter is a broad survey of markets, market participants, and some intermediate-term trends (monetary policy in a regime of abundant reserves; negative rates in Europe and Japan; and the changing nature of liquidity). Chapters 1 through 6 present the basic language and toolbox of the fixed income cosmos: arbitrage pricing; rates and spreads; DV01, duration, and convexity; and multi-factor and empirical hedging. Chapters 7 through 9 explain how term structure models are used for better understanding the shape of the term structure of interest rates; for pricing fixed income derivatives; and for relative value and even macro-style trading. Chapters 10 through 16 then delve into the details of several large and important markets: repurchase agreements or repo; note and bond futures; short-term rates and their derivatives; interest rate swaps; corporate bonds and credit default swaps; mortgages and mortgage-backed securities; and fixed income options.

While fixed income is an inherently quantitative subject, this book takes a very applied approach. All ideas are presented through examples, using market prices, events, or meaningful applications whenever possible. (A list of particularly extensive applications is given on the next page.) There is a lot of emphasis on orders of magnitude (e.g., "About how big is the interest rate swap market?" or "Approximately what is the DV01 or duration of a 10-year par Treasury?) and on fundamental concepts (e.g., "What does it mean for a position to be negatively convex?" or "What sorts of trades and positions have financing risk?).

This fourth edition is a comprehensive revision. All data on markets and market participants have been updated; all examples and applications updated; and all of the in-depth market presentations rewritten to reflect contemporary issues (e.g., variation margin of cleared interest rate swaps changed to "settled-to-market"). The timing of the edition is deliberate with respect to the transition away from LIBOR. While SOFR and other replacements are relatively young, it is time for a textbook treatment of these new reference rates and their associated derivatives.

We would like to thank Bill Falloon, who has enthusiastically supported this textbook for two decades, and Purvi Patel, for her patient and expert shepherding of this edition through the production process; Judy DiClemente, for her deeply thoughtful editing of the manuscript; Sienna Sihan Zhu for excellent and careful research assistance; Kristi Bennett for diligent copy editing; and the people who kindly and generously gave of their expertise and time to enrich the contents of this book: Viral Acharya, Amitabh Arora, Richard Cantor, Jonathan Cooper, Richard Haynes, David Lohuis, Lihong McPhail, Greg Perez, David Sayles, James Streit, and Regis Van Steenkiste.

EXTENDED EXAMPLES, APPLICATIONS, AND CASES

- Idiosyncratic Pricing of US Treasury STRIPS (Section 1.5)
- Relative Value Spreads of High-Coupon Treasuries (Section 3.6)
- P&L Attribution for an Outright Long in a High-Coupon Treasury (Section 3.8)
- Hedging a Century Bond (Sections 4.3 and 4.6)
- Hedging Stylized Pension Liabilities (Sections 4.8 and 5.4)
- Regression Hedge of the Johnson & Johnson 2.45s of 09/01/2060 with a 30-Year US Treasury (Section 6.1)
- Estimation of the Gauss+ Model over the Period January 2014 to January 2022 (Section 9.2)
- MF Global's Repo-to-Maturity Trades (Section 10.7)
- US Treasury Futures Basis Trades in March 2020 (Section 11.13)
- Extracting an Implied Path of Short-Term Rates from Fed Fund Futures, October 2021 *versus* January 2022 (Section 12.3)
- Tranche Structure of a CLO Issued in May 2019 (Section 14.1)
- Hertz CDS Settlement Auction, June 2020 (Section 14.11)
- The London Whale (Section 14.13)
- Three 30-Year FNMA Pools, 2018–2021 (Sections 15.5 through 15.7)
- Structure of a Credit Risk Transfer Security Issued in 2020 (Section 15.12)
- Pricing a Callable Bank of America Bond with Black-Scholes-Merton, August 2021 (Section 16.1)

List of Acronyms

ABS	Asset-backed securities
ADV	Average daily volume
AFX	American Financial Exchange
Ameribor	American Interbank Offered Rate
AP	Authorized participant
ARM	Adjustable-rate mortgage
ATM	At-the-money
AUD	Australian dollar
AXI	Across-the-Curve Credit Spread Index
BIS	Bank for International Settlements
BOJ	Bank of Japan
BRL	Brazilian real
BSBY	Bloomberg Short-Term Bank Yield Index
BSM	Black-Scholes-Merton
CAD	Canadian dollar
CCP	Central counterparty
CDS	Credit default swaps
CHF	Swiss franc
CLO	Collateralized loan obligation
CLOB	Central limit order book
CMO	Collateralized mortgage obligation
CMT	Constant-maturity Treasury
CP	Commercial paper
CPR	Conditional or constant prepayment rate
CRT	Credit risk transfer
CTD	Cheapest-to-deliver
CTM	Collateralized-to-market
CVA	Credit value adjustment
DTS	Duration times spread
DVP	Delivery versus payment
DV01	Dollar value of an ’01
EC	European Commission
ECB	European Central Bank
EFFR	Effective fed funds rate
ENNs	Entity-netted notionals

EONIA	Euro Overnight Index Average
ESG	Environmental, social, and governance
ESTER	Euro Short-Term Rate
ETF	Exchange-traded fund
EUR	Euro
EURIBOR	Euro Interbank Offered Rate
FCM	Futures commission merchant
FHLMC	Freddit Mac or Federal Home Loan Mortgage Corporation
FICC	Fixed Income Clearing Corporation
FNMA	Fannie Mae or Federal National Mortage Association
FOMC	Federal Open Market Committee
FRA	Forward rate agreement
FRN	Floating-rate note
FVA	Funding value adjustment
GBP	British pound
GC	General collateral
GCF	General collateral finance
GSE	Government-sponsored enterprise
HFT	high-frequency trading
HQLA	High-quality liquid assets
HQM	High-quality market-weighted
IDB	Interdealer broker
IM	Initial margin
IMM	Inside market midpoint
IMM	International money market
IO	Interest only
IOER	Interest on excess reserves
IORB	Interest on reserve balances
IRS	Interest rate swap
JPY	Japanese yen
LCH	London Clearing House
LCR	Liquidity coverage ratio
LIBOR	London Interbank Offered Rate
LTRO	Long-term refinancing operations
LTV	Loan-to-value
MAC	Market agreed coupon
MBS	Mortgage-backed securities
MPOR	Margin period of risk
MRO	Main refinancing operations
MTN	Medium-tern note
NAV	Net asset value
NOI	Net open interest
NPV	Net present value
NSFR	Net stable funding ratio

LSOC	Legally separated operationally comingled
OAD	Option-adjusted duration
OAS	Option-adjusted spread
OIS	Overnight index swap
ON	Overnight
OTC	Over-the-counter
OTR	On-the-run
PAC	Planned amortization class
PCA	Principal component analysis
PO	Principal only
PTF	Principal trading firm
P&L	Profit and loss
REMIC	Real estate mortgage investment conduit
RFQ	Request for quote
RRP	Reverse repo
RSAT	Replication synthetic asset transaction
RTM	repo-to-maturity
SARON	Swiss average rate overnight
SATO	Spread at origination
SDR	Swap data repositories
SEF	Swaps execution facility
SEK	Swedish krona
SEQ	Sequential pay class
SIMM	Standard initial margin model
SLUGs	State and local government series
SMM	Single monthly mortality
SOFR	Secured Overnight Financing Rate
SONIA	Sterling Overnight Interbank Average
STM	Settled-to-market
STRIPS	Separate Trading of Registered Interest and Principal of Securities
TAC	Targeted amortization class
TIBOR	Tokyo Interbank Offered Rate
TIPS	Treasury Inflation Protected Securities
TLTRO	Targeted long-term refinancing operations
TONAR	Tokyo Overnight Average Rate
UMBS	Uniform mortgage-backed securities
USD	United States dollar
VM	Variation margin
WAC	Weighted-average coupon
WALA	Weighted-average loan age
WAM	Weighted-average maturity

CHAPTER 0

Overview

0.1 GLOBAL FIXED INCOME MARKETS

Fixed income markets are large and global. Figure O.1 shows the outstanding amounts of debt securities, by residence of issuer. Debt securities are instruments designed to be traded, like bonds issued by corporations or by governments. Grouping by residence of issuer means, for example, that US Treasury bonds held by China's central bank are included in the total for the United States. As of March 2021, the global total of outstanding debt securities was about $123 trillion. For reference, the total capitalization of global equity markets at the time was $110 trillion, although stock market values are significantly more volatile.

Figure O.1 shows that the five largest issuers, in terms of amounts outstanding, are in the United States, the Eurozone, China, Japan, and the United Kingdom, which together comprise nearly 90% of the total. The Eurozone includes countries that both belong to the European Union (EU) and use the euro as a national currency. Some individual members of the Eurozone, indicated with asterisks in the figure, are significant issuers of debt securities on their own. Note that the figure displays their amounts outstanding with gray bars, but their contributions to the cumulative total are included once, with the Eurozone total.

Figure O.2 decomposes debt outstanding in the five largest regions by sector. The large fraction of government debt in Japan reflects decades of government borrowing and spending intended to stimulate the economy. The fraction of government debt in the United States, the Eurozone, and the United Kingdom is lower, at about 50%, but has increased significantly since the financial crisis of 2007–2009. Corporations in the United States are relatively more likely to issue bonds directly to the public, while corporations in the Eurozone, Japan, and the United Kingdom are relatively more likely to borrow funds from intermediaries, like banks, which, in turn, raise money from the public. While Figure O.2 includes the breakdown for debt in China, the relatively large role of the government in financial and nonfinancial enterprises makes comparisons across sectors and regions less meaningful.

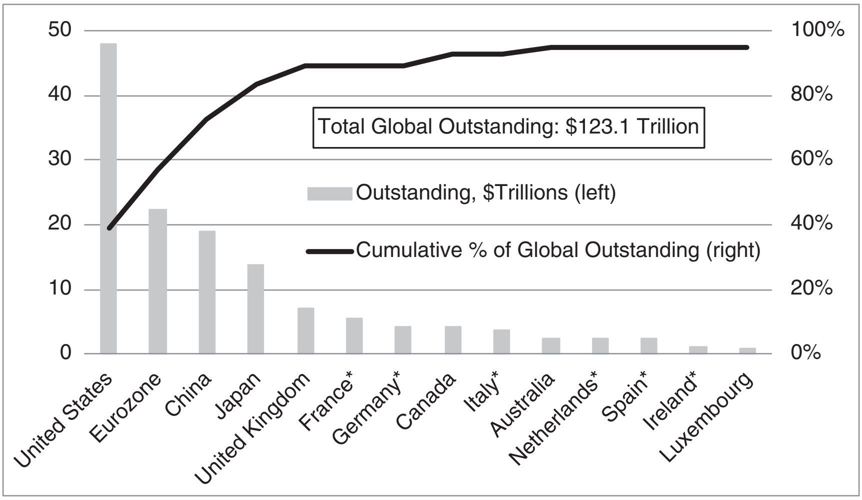

FIGURE 0.1 Global Debt Securities Outstanding, by Residence of Issuer, as of March 2021. Countries with an Asterisk Are in the Eurozone.
Sources: BIS; and Author Calculations.

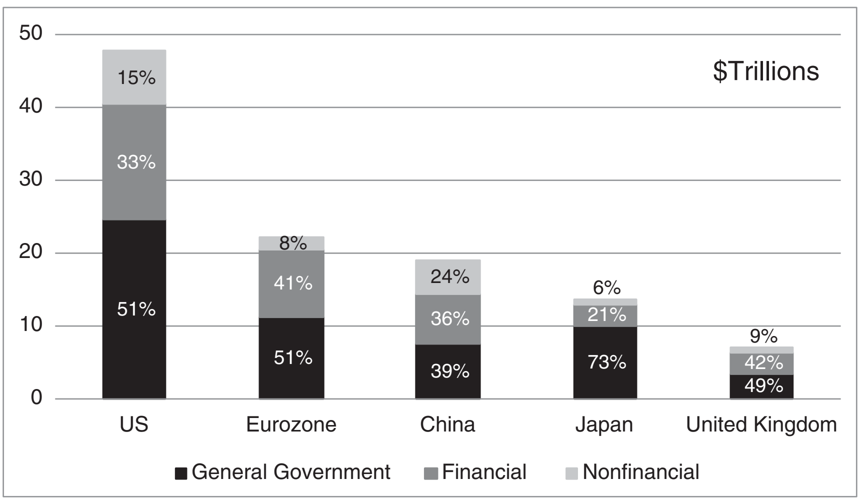

FIGURE 0.2 Global Debt Securities Outstanding, by Sector, as of March 2021.
Sources: BIS; and Author Calculations.

Table O.1 and Figure O.3 show the *notional* amounts of outstanding interest rate derivatives across the globe. These derivatives are described in later chapters, but derivatives essentially allow market participants to take positions on interest rates, whether for hedging, investment, or speculative purposes. The notional amount of a derivative is used to calculate the cash flows that one of the derivative's counterparties pays the other. Adding together all notional amounts, however, can significantly overstate market size. First, the largest market participants, namely dealers, tend to be simultaneously long and short nearly identical derivatives. Second, options are actually equivalent to only fractions of the notional amounts of their underlying securities. Later chapters elaborate on these points, but, for the purposes of this overview, this table and figure are reported in notional amounts.

While derivatives may trade in a particular locality, there is no sense in which derivatives are issued in one place or another: local regulations aside, any entity, residing anywhere, can enter into these derivatives contracts. A typical classification, therefore, is the currency in which the cash flows of the derivative are denominated. The first two columns of Table O.1 show notional amounts for swaps, options, and forward rate agreements. Most of the outstanding amounts are denominated in US Dollars (USD) and Euro (EUR). The quantities in the table hint at the overstatement of market size by notional amount: if the sizes of the markets for these USD derivatives were really $150 trillion, they would be larger than the combined size of all global debt securities markets. The second two columns of the table show the notional amounts of standardized, exchange-traded interest rate futures

TABLE O.1 Notional Amounts of Interest Rate Derivatives. Swaps, Options, and FRAs, as of June 2020; Futures and Futures Options, as of December 2020. Entries in $Trillions.

Swaps, Options, FRAs		Futures and Futures Options	
Currency	**Amount**	**Currency**	**Amount**
USD	152.1	USD	41.5
EUR	132.6	GBP	10.4
Other	67.1	EUR	9.5
GBP	54.3	BRL	1.6
JPY	37.1	CAD	1.0
CAD	14.3	AUD	0.9
SEK	5.3	Other	0.6
CHF	3.6		

USD: United States Dollar; EUR: Euro; GBP: British Pound; JPY: Japanese Yen; CAD: Canadian Dollar; SEK: Swedish Krona; CHF: Swiss Franc; BRL: Brazilian Real; AUD: Australian Dollar.
Source: BIS.

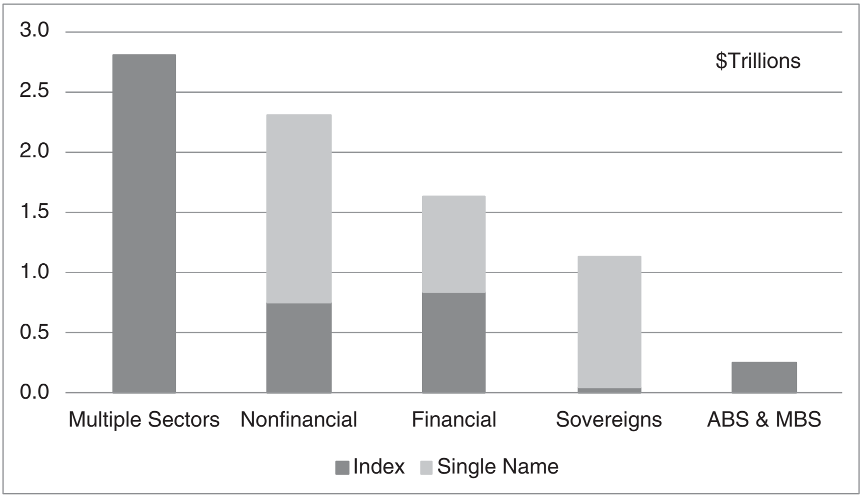

FIGURE 0.3 Credit Default Swaps, Notional Amounts Outstanding, by Sector and Type, as of June 2020.
Source: BIS.

and options. Amounts outstanding of USD-denominated contracts are by far the greatest, with those denominated in British Pounds (GBP) and EUR making up most of the rest of the overall market.

Finally, Figure O.3 gives the notional amount of credit default swaps (CDS) outstanding. These are discussed in detail in Chapter 14, but, roughly speaking, CDS allow investors to take positions that are equivalent to leveraged long or short positions in bonds with credit risk. The figure divides the market into credit sectors: CDS can be written on nonfinancial companies, financial companies, sovereigns, asset-backed securities (ABS), and mortgage-backed securities (MBS). Within each sector, a single-name CDS references a single credit (e.g., the government of Spain), while an index CDS references a portfolio of credits (e.g., 25 European financial companies). Note that the CDS market is much smaller in notional amount than the derivatives markets depicted in Table O.1.

0.2 US MARKETS

This section describes debt and loan instruments in the United States, categorized as in Figure O.4. The total amount outstanding across all instruments, as of June 2021, was $76.4 trillion.[1] By way of comparison, the market

[1]The data sources are different, but the inclusion of nontraded instruments here is the main discrepancy between this $76.4 trillion total and the $48 trillion total of

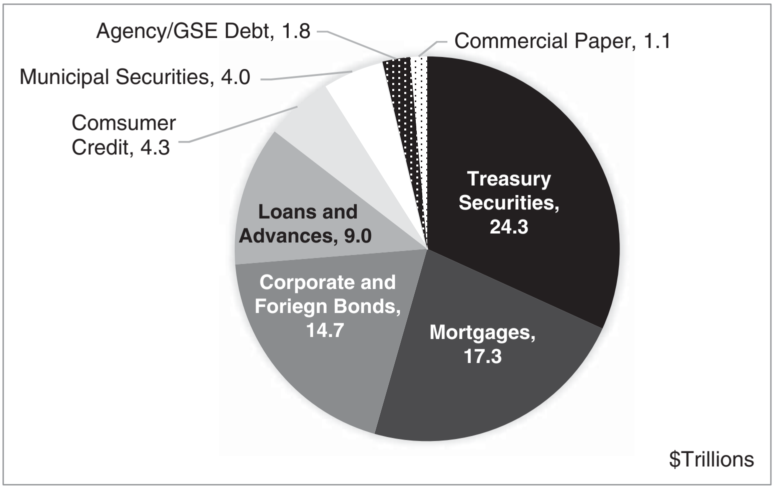

FIGURE O.4 Debt Securities and Loans in the United States, Amounts Outstanding, as of June 2021. GSE: Government-Sponsored Enterprise.
Sources: Financial Accounts of the United States, Board of Governors of the Federal Reserve System; and Author Calculations.

capitalization of US equities at the same time was about $45 trillion. Treasury securities and municipal securities are discussed in this section in some detail, while sectors discussed in later chapters of the book are treated very briefly here.

Treasury Securities

In less than a decade, Treasury securities have grown from the third largest category, behind mortgages and corporate and foreign bonds, to the largest category, at $24.3 trillion. When the US government spends more than it collects in taxes and fees, which has been the case for most of the last 50 years, it needs to borrow money to fund its deficit spending. It does so through the array of instruments shown in Figure O.5. Treasury *bills* or *T-bills* mature in one year or less and are *discount* securities, which means that they sell for less than, or at a discount from, their promised payment at maturity. Treasury *notes* and *bonds* are *coupon-bearing* securities; that is, they earn a fixed coupon or interest rate on their *principal*, *face*, or *par* amounts through maturity, and then repay that principal amount at

---

US debt securities in Figures O.1 and O.2. In particular, adding to that $48 trillion the $6.8 trillion of nonmarketable Treasury securities, $7.8 trillion of nonsecuritized mortgages, $9.0 trillion of loans and advances, and $4.3 trillion of consumer credit gives a total of about $76 trillion.

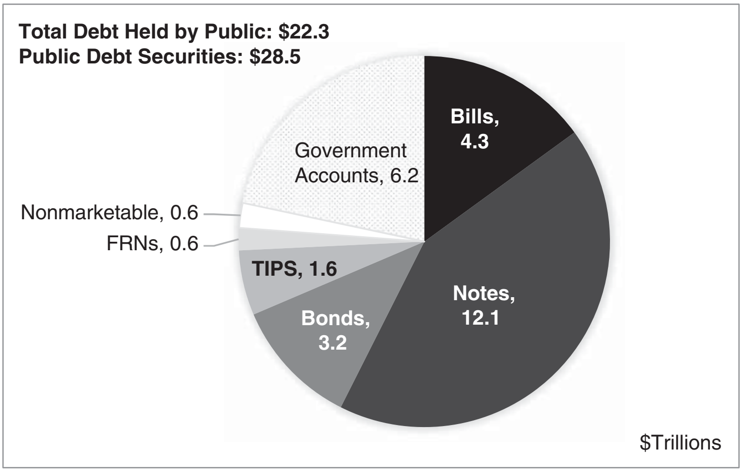

FIGURE 0.5 US Treasury Obligations, Amounts Outstanding, as of June 2021.
Source: US Treasury Bulletin.

maturity. Strictly speaking, and in the accounts of the government, notes are issued with 10 or fewer years to maturity, while bonds are issued with more than 10 years to maturity. The distinction had more meaning historically, when bonds were subject to a maximum, statutory rate of interest. In common parlance today, however, the words "notes" and "bonds" are used interchangeably. In any case, Chapter 1 describes the cash flows of Treasury notes and bonds in more detail.

Treasury Inflation Protected Securities (*TIPS*) protect investors against inflation with principal amounts that increase or decrease with changes in the consumer price index (CPI). Consider, for example, a TIPS with a principal amount of $100 and a coupon rate of 1% per year. If CPI has increased by 10%, the principal amount of the TIPS will have increased to $110, and the investor earns 1% on that higher principal amount. This investor is just as well off having $100 and earning $1 per year at the original price level as having $110 and earning 1%× $110 or $1.10 per year at a price level that is 10% higher. Hence, TIPS earn a fixed *real* or inflation-adjusted return, while coupon-bearing Treasuries earn a fixed *nominal* or dollar return.[2] For this reason, by the way, in discussions that include both TIPS and Treasury bonds, the latter are often referred to as *nominal bonds*. In any case, while comprising only 5.6% of the total in

[2]Over a period of deflation, TIPS principal amount may fall below original principal for the purposes of calculating interest, but, at maturity, TIPS return at least the original principal amount.

Figure O.5, TIPS have an outsized importance as measuring the market's perception of inflation and the price of inflation risk. As of January 2022, for example, the rate on five-year nominal Treasury bonds was about 2.8% higher than the rate on five-year TIPS. Roughly expressed, therefore, the market expects an average inflation rate of 2.8% over the subsequent five years. More precisely expressed, given inflation expectations and risk preferences, investors require a premium of 2.8% to buy five-year nominal bonds and to assume inflation risk over that horizon.

Figure O.5 next lists *floating-rate notes* (FRNs). These are relatively new, having been first issued in January 2014. FRNs are sold with two years to maturity, and they pay a variable rate of interest equal to the going rate on 13-week T-bills plus a fixed spread. This spread entices some investors, who might otherwise roll investments of short-term T-bills, to sacrifice some liquidity and buy two-year FRNs instead. From the perspective of the Treasury, FRNs lock in funding for two years, but at a cost only slightly above that of short-term bills.[3] In fact, over 2021, FRNs were sold at a spread of less than five basis points.[4] In addition, somewhat cynically, FRNs seem to cost less than two-year notes, because government accounting of interest rate cost does not penalize the risk of rates increasing in the future. In any case, the issuance of FRNs has remained limited, comprising about 2% of the total in Figure O.5.

US Treasury issues are among the most actively traded securities in the world. This is due, in good part, to the significant role of the US dollar in international markets and to the perception of US government debt as among the best stores of value available. An additional explanation, however, is the careful management of Treasury debt issuance. More specifically, the US Treasury sets a regular *auction schedule* that lets investors know, in advance, which securities will be sold when and in what quantities. Furthermore, the Treasury has gradually modulated this schedule over many years to suit both the borrowing needs of the government and changing market conditions. To appreciate these points, Table O.2 describes the auction schedule as of January 2022.

"Issue frequency" describes how often new securities of each type are issued. The Treasury has settled, for example, on issuing new notes with two, three, five, and seven years to maturity every month, while issuing new

[3]While there would seem to be no particular need to lock in longer-term funding, because the US Treasury has never had a problem rolling over its short-term borrowings, prudent debt management avoids scenarios in which too much debt matures and needs to be refinanced over a relatively short period of time.

[4]A basis point is 0.01%. The difference between a rate of 1.01% and 1.00% is one basis point, and a rate of 1% might be referred to as a rate of 100 basis points. The spread of five basis points in the text, therefore, is a spread of 0.05%.

TABLE O.2 US Treasury Auction Schedule, as of January 2022.

Security	Issue Frequency	Reopenings
4-, 8-, 13-, and 26-week bills	Weekly	
52-week bills	Every 4 weeks	
2-, 3-, 5-, and 7-year notes	Monthly	
10-, 20- and 30-year notes/bonds	Quarterly	Monthly
5-year TIPS	Semiannually	2 months after issuance
10-year TIPS	Semiannually	Every 2 months
30-year TIPS	Annually	Semiannually
2-year FRNs	Quarterly	Monthly

TIPS: Treasury Inflation Protected Securities; FRNs: Floating Rate Notes.
Source: US Department of the Treasury.

10-year notes, along with new 20- and 30-year bonds, every quarter. Both the set of issues and their frequency have changed over time, however. For example, issuance of the 20-year bond was eliminated in 1986 and brought back in May 2020, while issuance of 30-year bonds was stopped after August 2001 and resumed in February 2006. And in 2000, the US Treasury introduced the concept of *reopenings*, which means auctioning or selling more of an existing issue. For example, Table O.2 reports that 10-year notes are issued quarterly and reopened monthly. In August 2021, the Treasury sold about $59 billion of a new 1.25% 10-year note, that is, a note that had not been issued before, that pays interest on principal at an annual rate of 1.25%, and that matures on August 15, 2031. The following month, in September 2021, the Treasury sold another $42 billion of that same issue, that is, more notes with a coupon of 1.25% that mature on August 15, 2031. And again the next month, in October 2021, the Treasury sold another $41 billion of that same issue. In November 2021, however, a quarter after the first issuance of the 1.25% 10-year notes, the Treasury sold $62 billion of a new 10-year note, with a coupon of 1.365% and a maturity date of November 15, 2031.

One result of the auction schedule is that the most recently issued bonds of each type and maturity tend to be the ones most actively traded. In the previous example, as of November 15, 2021, the just-issued 1.365% notes maturing on November 15, 2031, are called the 10-year *on-the-run* notes and are likely to become the most actively traded of notes with approximately 10 years to maturity. The 1.25% notes maturing on August 15, 2031, which had been the 10-year on-the-run notes, become the *old* notes, and over time, as even newer 10-year notes are issued, become the *double-old* notes, the *triple-old* notes, etc.

Returning to Figure O.5, *nonmarketable* securities is another small category of Treasury issuance. Included in this category are about $140 billion of savings bonds, which are discount securities sold directly to individual investors, and about $120 billion of *State and Local Government Series* bonds, commonly known as *SLUGs*, which are mentioned later in the context of municipal bonds.

The final category in Figure O.5 refers to Treasury bonds sold into *government accounts*. These bonds are also nonmarketable and represent debt that the US government owes itself. The social security trust funds, for example, at the end of 2020, held about $3 trillion in Treasury bonds. Most people argue that these bonds do not represent any additional Treasury indebtedness, or, in other words, that there is no difference between social security benefits being paid by the Treasury directly or being paid indirectly through the payment of interest and principal on bonds in the social security trust funds. By this logic, Treasury bonds held in government accounts are excluded from most descriptions of government indebtedness. More specifically, in terms of Figure O.5, US government debt is usually equated to "Total Debt Held by the Public," or $22.3 trillion, rather than to "Public Debt Securities," which adds the $6.2 trillion held in government accounts, for a grand total of $28.5 trillion.

In discussing the magnitude of government debt, and in comparing government indebtedness across countries, debt held by the public is usually normalized by *gross domestic product* (GDP), which measures the value of the goods and services produced in a country in a single year. The idea here is that countries with greater GDPs can safely carry greater levels of debt. With US GDP at about $22 trillion as of June 2021, the ratio of debt to GDP in the United States is about 100%, which is extremely high by historical standards. The ratio was over 100% during World War II; subsequently declined to between 20% and 50%; climbed to about 80% in the aftermath of the financial crisis of 2007–2009; and then shot up to above 100% in the wake of the COVID pandemic and economic shutdowns. For comparison purposes, the ratio of debt to GDP in Japan is currently over 230%; in Greece about 175%; in France about 100%; in the United Kingdom about 85%; in Germany and China, less than 60%; and in Switzerland, less than 40%.

Despite the historically high ratio of US debt to GDP, foreign investors continue to find US Treasuries attractive and hold a large fraction of the amount outstanding. As of June 2021, investors outside the United States held $7.2 trillion Treasuries, or 33% of the $21.8 trillion marketable securities shown in Figure O.5. These holdings include $1.28 trillion (5.9%) in Japan, $1.06 trillion in China (4.9%), and $0.53 trillion (2.4%) in the United Kingdom.

Municipal Securities

The $4 trillion municipal securities market includes more than 50,000 issuers and approximately one million individual bond issues. Municipal bonds, sometimes called *municipals* or simply *munis*, are issued by states and local governments to fund their expenditures. Unlike most fixed income markets, the muni market is dominated by retail investors: as of June 2021, over 70% of principal outstanding was held directly by individuals, or indirectly by individuals through mutual funds and other investment vehicles.

Interest payments from munis are exempt from federal tax so long as funds raised from selling those munis are used for public projects. Therefore, investors who pay federal taxes are willing to accept lower rates of interest from munis than they would from bonds whose interest is taxed at the federal level, like corporates and Treasuries, controlling, of course, for differences in credit quality. And municipal issuers can raise funds at rates below what they would otherwise have to pay. The story is somewhat more complicated, however, because capital gains on price appreciation from munis is not exempt from federal tax. To avoid this tax, however, the market has evolved to minimize the proportion of return in the form of (taxed) price appreciation rather than (untaxed) interest. More specifically, most munis – in the current low-rate environment – are issued at a coupon rate of 5% and, consequently, at a premium to par, that is, at a price greater than face amount. In this way, these munis are unlikely to trade at a discount and thus subject subsequent buy-and-hold purchasers to taxes on capital gains. With respect to state taxes, the treatment of municipal interest varies by state. Most states tax bonds issued in other states and exempt their own bonds, while some states tax both their own and other states' bonds. Washington, D.C., does not tax any municipal bond interest, and Utah exempts municipals issued in states that exempt Utah's bonds! Utah's exemption is more significant than it may seem, because states with no income tax automatically qualify. The final piece of the tax story is that about $500 billion of municipals do not qualify for the federal tax exemption, because their proceeds are used for working capital, for funding private business development, or for refinancing existing debt through particular kinds of transactions.[5] These *taxable munis* pay interest at rates comparable to those in the corporate and Treasury markets, again controlling for differences in credit quality.

Muni bonds can be divided into three broad groups. *General obligation (GO) bonds*, which constitute about 25% of the market, are backed by the taxing power of the issuing municipality. *Revenue bonds*, which constitute about two thirds of the market, are backed by revenues from particular

[5]Historically, municipalities had been allowed to refinance existing tax-exempt debt in advance of maturity or any call option with the proceeds of new, tax-exempt debt, but this practice was outlawed in 2017. For a full treatment of this and other related issues, see Kalotay, A. (2021), *Interest Rate Risk Management of Municipal Bonds*, Andrew Kalotay Associates, Inc.

projects, like tolls from a bridge or highway. And within revenue bonds is a $600 billion subset of *industrial revenue bonds*, through which municipalities issue (tax-exempt) debt to raise funds for private enterprises engaged in qualifying projects. The third group consists of *prerefunded* or *defeased* bonds. A muni in this category has been canceled as a municipal liability by the setting aside of sufficient cash and Treasury securities to fund all of its remaining payment obligations. These prerefunded or defeased bonds continue to exist and trade, but with their credit risk improved to that of Treasuries.[6]

Default rates on municipal securities have been very low. From 1970 to 2020, Moody's reports a five-year cumulative default rate for the sector of 0.08%, which is much lower than the equivalent for corporate bonds, as discussed in Chapter 14. Nevertheless, credit risk is an important consideration for investors in the muni market, with underfunded pension obligations a particularly significant and perennial cause for concern. Historically, GO bonds, which are backed by the issuer's taxing power, were perceived as safer than revenue bonds, which are backed by particular sources of revenue that could diminish or disappear over time. This perception changed significantly, however, with the bankruptcy of Detroit in the summer of 2013. The settlement provisions that emerged in November 2014 were a combination of law, political forces, and negotiations across many interested parties. The results can be roughly summarized as follows. First, while a court ruled that state laws protecting pension benefits were trumped by federal bankruptcy law, Detroit pensions recovered over 95% of their value, although cost of living adjustments were reduced or eliminated. Health and life insurance benefits, however, recovered only 10% of value. Second, holders of water and sewer bonds, which had strong legal claims to the associated revenue streams, suffered no loss of principal. Third, GO bonds suffered significant losses, depending on their exact provisions. *Unlimited tax* bonds carry a pledge by issuers to raise taxes, if necessary, to pay bondholders. The unlimited tax bonds involved in the Detroit bankruptcy were, in addition, backed by specific, segregated, and voter-approved tax receipts. Nevertheless, these bonds recovered only 74% of principal, and this negotiated, less-than-full recovery was attributed to the dwindling of their dedicated tax receipts due to deteriorating economic

[6]When a municipality defeases a tax-exempt bond issue, it needs to purchase Treasury securities, but is prohibited – to prevent tax arbitrage at the expense of the federal government – from investing at a higher rate than the rate on the defeased, tax-exempt issue. This problem is solved by purchasing SLUGs directly from the Treasury, which are mentioned earlier in the chapter. SLUGs are the same credit quality as other Treasuries but pay a lower interest rate to suit the needs of the particular defeasance.

conditions in Detroit. Worse off, however, were Detroit's *limited tax* GO bonds, which had neither a pledge of unlimited tax increases nor dedicated tax receipts. They recovered only 34% of principal. The revelation that GO bonds can be treated like unsecured claims changed credit analysis and pricing in the muni market.

Before the financial crisis of 2007–2009, a majority of muni bonds were insured against default, for a fee, of course, by private insurers. In fact, most of Detroit's GO bonds were insured, which meant that insurers, rather than investors, suffered the losses described in the previous paragraph. In any case, muni insurers suffered massive losses during the financial crisis, not from their muni businesses, but from having insured mortgage-related products. The ensuing damage to the industry ultimately resulted in less than 5% of new muni issues being insured. More recently, perhaps in part due to the Detroit bankruptcy, and perhaps in part due to the COVID pandemic and shutdowns, there has been a resurgence of muni insurance, rising to perhaps 10% of new issues.[7]

Other US Markets

Mortgages. The second largest sector in Figure O.4 contains mortgages, at $17.3 trillion. Mortgage loans are used to purchase properties and are collateralized by those same properties. Mortgage balances finance the purchase of one- to four-family residences (70%); commercial property (18%); multi-family residences (10%); and farms (2%). A remarkable feature about the US mortgage market is that only about 45% of mortgage loan balances are held by the original lenders. The remaining 55% of balances are *securitized*, that is, sold by the original lenders; packaged into securities; and then sold to investors. Chapter 15 describes this market in much greater detail.[8]

Corporate and Foreign Bonds. The third largest sector in Figure O.4 comprises corporate and foreign bonds, at $14.7 trillion, which includes bonds sold by US nonfinancial corporations (45%); by US financial corporations (31%); and by foreign corporations to US investors (24%), all to raise money to fund their operations and corporate transactions. Chapter 14 describes this market in detail.

[7] See Gillers (2020), "Bond Insurance Returns to the Muni Market in a Big Way," *The Wall Street Journal*, October 22; and Moran, D. (2020), "Municipal Bond Insurance Busier Than Ever After Decade-Long Slump," *Insurance Journal*, June 26.

[8] Because many of the MBS issued by the Federal National Mortgage Association (FNMA) and the Federal Home Loan Mortgage Corporation (FHLMC) have been consolidated on to their balance sheets, the Financial Accounts of the United States include these as part of the debt of government-sponsored enterprises (GSEs). Figure O.4 and the numbers in this paragraph exclude these MBS from Agency/GSE debt so as not to double count the underlying mortgages.

Loans and Advances. Continuing counterclockwise in Figure O.4, a little less than half of the $9.0 trillion of Loans and Advances are made by banks, and the rest by an assortment of nonfinancial and financial entities. An important feature of this sector is the securitization and trading of bank loans, which is described in Chapter 14. Consumer credit, at $4 trillion, is discussed in the next section, in the context of household balance sheets.

Agency/GSE Debt. The penultimate category of instruments in Figure O.4 includes "agencies" or "agency bonds," which are issued by government agencies and GSEs. These entities span a range of associations with the federal government, and their debt issues enjoy varying levels of support from the federal government. At one extreme are the Federal Housing Administration (FHA), the Small Business Administration (SBA), and the Government National Mortgage Association (GNMA). These agencies are part of the government, and their debt issues are backed by the "full faith and credit" of the United States. Moving from the full-faith-and-credit framework, the Tennessee Valley Authority (TVA), which provides electricity for local power companies in Tennessee and surrounding states, is considered a federal agency. Formally, however, it is a corporation that is wholly owned by the US government, and its debt is backed by its own revenues, not – at least explicitly – by the federal government. Further away from full-faith-and-credit are the Federal National Mortgage Association (FNMA) and the Federal Home Loan Mortgage Corporation (FHLMC). They are known as GSEs, but are owned by private shareholders and, having failed during the financial crisis of 2007–2009, are now under government conservatorship. Leading up to the crisis, their debt was not explicitly backed by the government, but the market expectations for full government support of their bonds was fully realized during and after the crisis. These two entities are discussed in great detail in Chapter 15.

Commercial Paper. The final category in Figure O.4 is Commercial Paper (CP), through which the most highly rated corporations borrow short-term funds from the public. CP is discussed further in Chapter 14.

0.3 US MARKET PARTICIPANTS

This section describes sectors of market participants. The volumes of debt securities and loans that appear as assets on the financial balance sheets of various sectors are shown in Figure O.6, while the volumes that appear as liabilities are shown in Figure O.7. Households, nonfinancial business, and general government (i.e., federal and municipal) hold a combined $10.0 trillion of debt securities and loans for their own savings or cash management purposes. Banks, various fund vehicles, insurance companies, pension funds, and other financial entities hold debt securities and loans as intermediaries, investing and managing funds for the ultimate benefit of

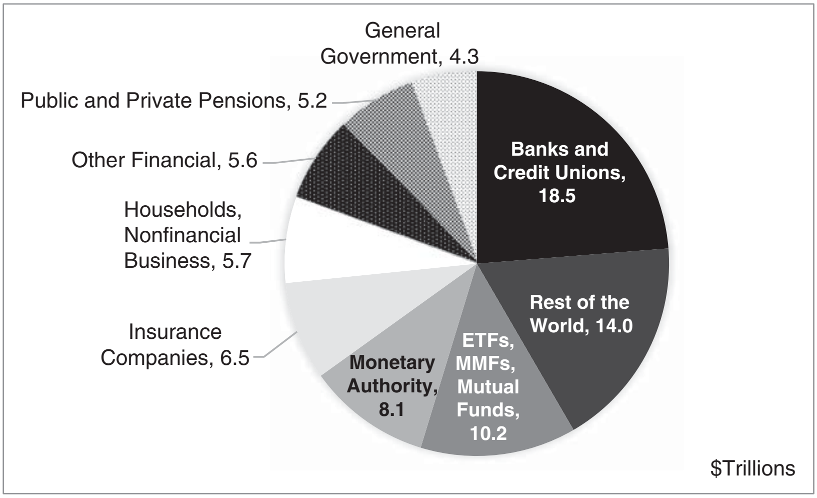

FIGURE 0.6 Financial Assets of Various Sectors, as of June 2021. ETF: Exchange-Traded Fund; MMF: Money Market Fund.
Sources: Financial Accounts of the United States, Board of Governors of the Federal Reserve System; and Author Calculations.

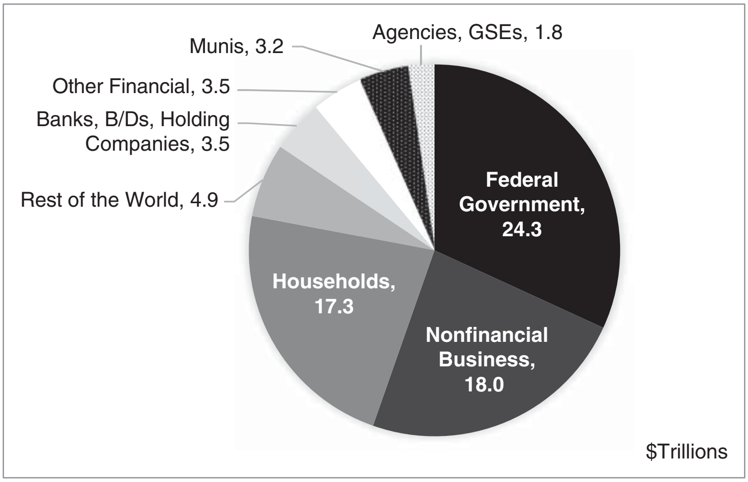

FIGURE 0.7 Financial Liabilities of Various Sectors, as of June 2021. B/D: Broker-Dealer.
Sources: Financial Accounts of the United States, Board of Governors of the Federal Reserve System; and Author Calculations.

others. The Monetary Authority, that is, the Federal Reserve, holds assets in the implementation of monetary policy. And the rest of the world, or foreign entities, invest $14.0 trillion in US debt securities and loans on their own or as intermediaries. On the liability side, the federal government borrows $24.3 trillion, and municipalities borrow $3.2 trillion to finance government expenses that are not covered by taxes and fees. Households borrow $17.3 trillion mostly to finance consumption, and nonfinancial businesses borrow $18.0 trillion to finance investments and acquisitions. The remaining financial sectors borrow as intermediaries, using funds borrowed from some sectors to lend or invest elsewhere. And finally, the rest of the world borrows $4.9 trillion through US debt securities and loans. The text now discusses several of these sectors in greater detail.

Households

Table O.3 gives the financial assets and liabilities of the household sector.[9] The financial assets of the sector far exceed its financial liabilities,

TABLE O.3 Financial Assets and Liabilities of Households, as of June 2021. Amounts are in $Trillions.

Instrument	Amount	Percent of Total
Total Financial Assets	113.1	100
Pension Entitlements	31.0	27.4
Corporate Equities	30.5	26.9
Deposits and MMFs	17.1	15.1
Equity in Noncorporate Business	13.7	12.1
Mutual Funds	12.3	10.9
Debt Securities and Loans	5.1	4.5
Life Insurance Reserves	1.9	1.7
Other	1.7	1.5
Total Liabilities	17.7	100
Home Mortgages	11.3	63.8
Consumer Credit	4.3	24.1
Other	2.1	12.1

Sources: Financial Accounts of the United States, Board of Governors of the Federal Reserve System; and Author Calculations.

[9]The household sector, in this data set, includes private equity funds, domestic hedge funds, personal trusts, and nonprofits, with combined assets that could easily exceed $10 trillion. Their inclusion here results in some overstatement of the extent to which individuals invest directly in equity and debt markets.

meaning that the sector, as a whole, has significant net worth. The sector invests significant amounts directly in equity and debt markets, although a greater fraction of assets is invested through intermediaries, like pensions, mutual funds, and the savings components of life insurance policies. A large fraction of financial assets is also held in near cash equivalents, that is, in deposits and shares of money market funds. On the liability side, households borrow mostly through mortgages, which are the subject of Chapter 15, or through consumer credit.

The primary components of consumer credit are home equity loans, auto loans, credit card loans, and student loans. The outstanding credit balances of these sectors over time are depicted in Figure O.8. Home equity loan balances grew rapidly with the run-up of housing prices preceding the financial crisis of 2007–2009 and have declined steadily since.[10] Credit card loans also grew going into the crisis, though not as dramatically, and also declined after the crisis, but had then recovered, until declining again with the COVID pandemic and shutdowns. Auto loans, while also declining through the crisis, seemingly emerged as taking the place of declining home equity loans and relatively flat credit card loans. Student loans have increased independently of the economic cycle, which is a policy result: the federal government has

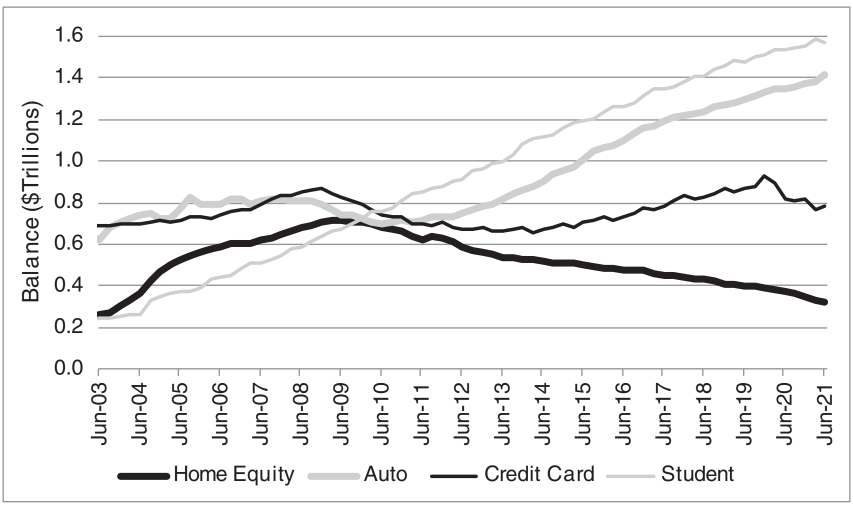

FIGURE O.8 Balances of Consumer Credit Sectors.
Source: Quarterly Report on Household Debt and Credit, Federal Reserve Bank of New York.

[10]Home equity loans are second liens, which are subordinate to first home mortgages with respect to creditor priority.

been guaranteeing student loans for decades and, since 2010, directly owns all of its student loans. As of March 2021, government student loan balances comprised over 92% of the total.[11]

Figure O.9 reports balances that are 90 or more days *delinquent*, that is, balances on which the borrower has not made payments for 90 or more days. Not surprisingly, home equity delinquencies increased during and for several years after the financial crisis, as falling housing prices made it impossible for many homeowners to recover their outstanding mortgage and loan balances by selling their homes. Home equity delinquencies have since steadily declined, perhaps reflecting more careful underwriting in the aftermath of the crisis. Credit card delinquencies also increased after the crisis, perhaps reflecting weakened consumer balance sheets, but have since declined to pre-crisis levels. Auto loan delinquencies increased after the crisis as well, then fell, and increased again, perhaps an expected consequence of the rapid growth in balances. Student loan delinquencies have been increasing for many years, which has raised a number of policy concerns, from the perspective of both students and the federal government. Note that the precipitous decline in student loan delinquencies in the second half of 2020 is an artifact of federal COVID forbearance programs.

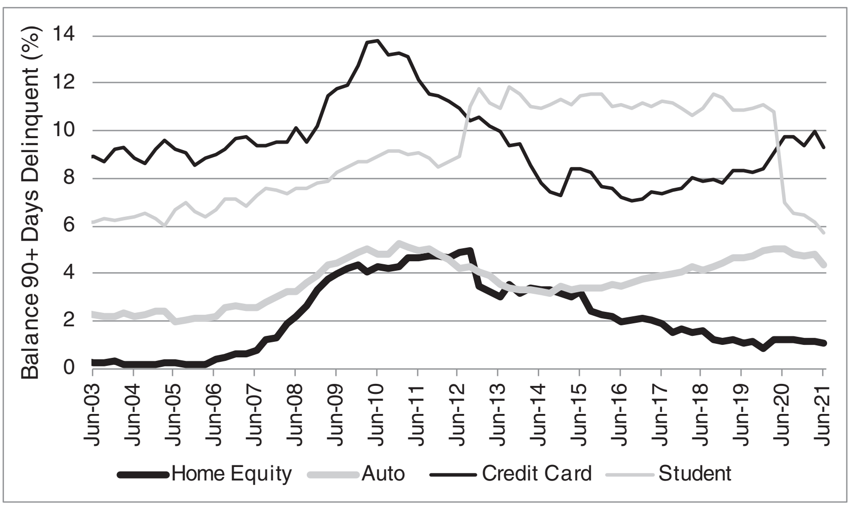

FIGURE O.9 Delinquencies in Consumer Credit Sectors.
Source: Quarterly Report on Household Debt and Credit, Federal Reserve Bank of New York.

[11]See, for example, Amir, E., Teslow, J., and Borders, C. (2021), "The MeasureOne Private Student Loan Report," June 15.

Nonfinancial Business

Figure O.10 shows the composition of liabilities for nonfinancial businesses, separated into corporate businesses, which are likely to be relatively large, and noncorporate businesses, which are likely to be relatively small. Corporate businesses, which are more likely to have track records of earnings and established creditworthiness, have broader access to markets. They can sell debt securities to raise 30% of their financial liabilities, while noncorporate businesses sell essentially none. The "Other" source for corporate businesses includes direct investment from abroad, which is not at all a part of small business liabilities. Noncorporate businesses, then, rely mostly on mortgages, for which they need only unencumbered property. The differences in the liability structures of these two groups reflect the life cycle of business borrowings, from "family and friends," to bank loans, to investor groups, to private placements of debt, and finally – for the largest and most established companies – to public debt securities. Chapter 14 discusses private placements and public debt issues in more detail.

Commercial Banks

Commercial banks accept deposits from their customers, pay them interest, and offer them safety – in part through federal deposit insurance – and immediacy or liquidity, that is, the ability to withdraw funds whenever necessary. Deposits, which are not classified as debt securities or as loans, are not

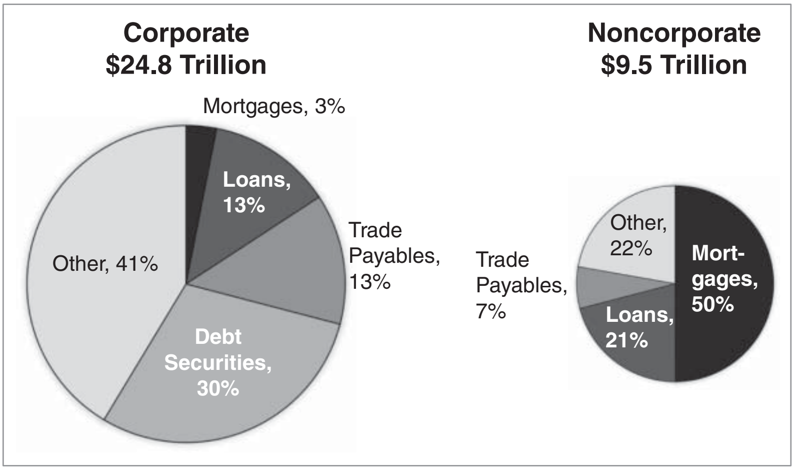

FIGURE O.10 Nonfinancial Business Liabilities, Corporate and Noncorporate, as of June 2021.
Source: Financial Accounts of the United States, Board of Governors of the Federal Reserve System.

included in Figure O.7, but constituted 93% of commercial bank liabilities. Other liabilities include investments by the bank's parent company, long-term debt, commercial paper, assorted loans, and *repurchase agreements*, or *repo*, which are loans collateralized by debt securities, as discussed in Chapter 10.

While deposits are certainly the main source of funding for bank investments, they are actually a product or output of banking, just like business loans and mortgages are bank products. Depositors value the safety and immediacy of deposits, and they incorporate deposits into their management of cash and liquidity. Furthermore, banks actively manage the liquidity they offer to depositors, in part by having some fraction of liabilities in longer-term debt, and in part by investing some fraction of assets in liquid products that can be sold quickly and easily to meet any unexpected withdrawals of deposits.

Turning to assets, Figure O.11 shows the asset composition of large and small commercial banks. Commercial and industrial (C&I) loans, real estate loans, and consumer loans are all considered the main business of banking. In addition to these assets, however, along the lines of the previous paragraph, banks hold liquid assets. The most liquid, of course, are cash, reserves or deposits at the Federal Reserve, and other money market (MM) instruments. But these assets have the disadvantage of earning very low rates of return. Therefore, to earn higher rates of return while maintaining satisfactory liquidity profiles, banks also hold Treasuries, agency securities, and MBS, which can be sold relatively easily should the need arise.

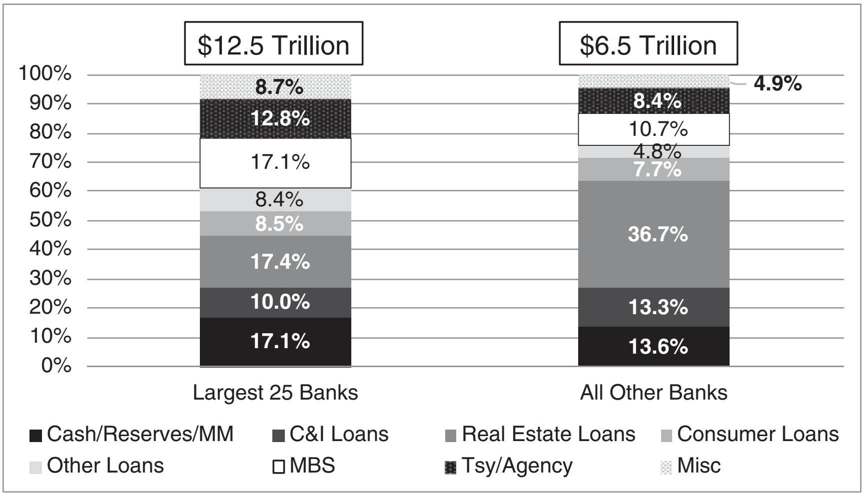

FIGURE O.11 Assets of Commercial Banks, Largest 25 Banks and All Other Banks, as of June 2021.
Source: Assets and Liabilities of Commercial Banks in the United States, Board of Governors of the Federal Reserve System.

Figure O.11 also reveals some differences between large and small banks. First, commercial bank assets are highly concentrated. There are over 4,000 commercial banks in the United States, but $12.5 trillion of the sector's $19 trillion of assets, or 66%, are held by the largest 25 banks.[12] As an aside, the number of banks in the country has been declining gradually but swiftly: there were over 14,000 banks in 1984. This decline is likely an adjustment from historical restrictions on interstate banking and branching that prevented larger banks from satisfying market demand. In any case, a second difference between the largest and smaller banks is the difference in the fractions of their assets in real estate loans: 17.4% for the largest banks and 36.7% for the smaller banks. The concentration of a small bank's assets in real estate loans, which are often local, can challenge the bank's viability through regional economic downturns.

Life Insurance Companies

Life insurance products often pay death benefits, of course, but they are also often savings vehicles, through which policy holders invest funds with the advantages of tax deferral. Life insurance companies are, therefore, financial intermediaries that collect and invest premiums so as to meet their obligations under policies sold and to earn additional returns for their shareholders. Furthermore, long-term fixed income assets are natural hedges to the long-term nature of their policy liabilities. Reflecting these considerations, the asset portfolios of life insurance companies contain large fractions of corporate bonds and equities, 36.4% and 8.5%, respectively, with an additional 18.8% in mutual fund shares that are some mix of bonds and equities. In fact, their corporate bond investments make life insurers very significant players in that market: their direct holdings of $3.5 trillion of corporate bonds comprise about 24% of the total $14.7 trillion outstanding. While Treasuries are theoretically useful as a match for long-term liabilities, they do not earn enough to meet insurer return hurdles. As a result, Treasuries comprise only 2.4% of life insurance company assets. Finally, life insurance companies also use derivatives to achieve their return and hedging objectives.

Pension Funds

Historically, the majority of pensions were *defined benefit (DB)* plans, in which the sponsor promises to pay retirees according to a formula that depends on the number of years worked, contributions to the plan, salary

[12] With respect to competitiveness, note that concentration within the banking sector does not, on its own, account for competition with nonbanking entities, like money market funds, nonbank mortgage lenders, and some fintech firms.

history, etc. Sponsors collect employee and their own contributions into a pension fund, and then invest the assets of that fund so as to be able to honor promised obligations. A low-risk strategy combines relatively high contributions to the pension fund with low-risk investments, while a high-risk strategy combines relatively low contributions with aggressive investments. In any case, investment risk in DB plans resides with the sponsor, which, in the end, is responsible for paying the promised benefits. For decades now, however, *defined contribution (DC)* plans have become more important. In these plans, employees and employers contribute to individual employee accounts, and each employee can typically choose among a few investment options. Upon retirement, employee benefits are determined completely by the funds accumulated in their respective accounts. Hence, in DC plans, investment risk resides with employees. Employers can, of course, offer both types of plans or a hybrid of the two types.

Government employees, at the federal, state, and local levels, nearly always have DB plans or an option to participate in a DB plan. In the private sector, however, the trend has been for corporations and other employers to avoid the risks and costs of managing pension funds, that is, to migrate from DB to DC plans. In 1975, there were about 33.0 million participants in private DB plans and 11.5 million in private DC plans. In 2019, the numbers were 32.8 million and 109.1 million, respectively.[13] Furthermore, corporations are actively shedding DB pension fund risk through *pension risk transfer* transactions, in which they pay insurance companies to assume their pension liabilities. In any case, for any portfolio manager of a DB plan, long-term debt instruments naturally hedge the present value of fixed liabilities. On the other hand, allocations to equities hold out the prospect of having to make smaller contributions to the fund. A study of the largest 100 DB plans offered by US public companies in 2020 found that 50% of assets were invested in fixed income, 32% in equity, and 18% in other categories (e.g., real estate, private equity, hedge funds).[14]

Money Market Funds

As mentioned earlier, in the context of deposits, there exists significant demand for assets that provide both safety and immediacy. Money market funds, created in the 1970s, were designed to offer safety and immediacy, while paying higher rates than banks at the time were allowed to pay on deposits. Money market funds are divided into three broad categories:

[13]Employee Benefits Security Administration (2021), "Private Pension Plan Bulletin Historical Tables and Graphs, 1975–2019," September.
[14]Wadia, Z., Perry, A., and Clark, C. (2021), "2021 Corporate Pension Study," Milliman White Paper, April.

government funds, which purchase only short-term, government-backed debt; prime funds, which invest predominantly in short-term, high-quality corporate debt, like commercial paper; and tax-exempt funds, which invest in short-term, high-quality, tax-exempt municipal debt. Before changes implemented after the financial crisis of 2007–2009, investors could buy money market fund shares for $1 per share; their money was invested in relatively safe and liquid assets; and, except in extraordinary circumstances, they could sell their shares at any time for $1 per share. More specifically, a fund that i) complied with Securities and Exchange Commission (SEC) rules governing the safety and liquidity of fund investments, known as 2a-7 rules; and ii) had a portfolio or net asset value (NAV) corresponding to a value per share of between 99.5 cents and $1.005, could offer and redeem shares at a "fixed NAV" or "stable NAV" of $1 per share. But if the value of the fund fell such that the value per share fell below 99.5 cents, the fund would "break the buck" and shares would no longer be redeemed at $1, but rather at a value corresponding to the fund's NAV. Hence, money market fund shares were very similar to bank deposits, but did not have the benefit of an explicit government guarantee, like federal deposit insurance. Instead, money market shareholders had to rely on their fund sponsors or management companies to make up for any NAV shortfalls. Only one money market fund had ever broken the buck, in 1994, but it was to happen a second time during the financial crisis of 2007–2009.

In September 2008, a few days after the failure of the mortgage government-sponsored enterprises, FNMA and FHLMC, and a few days before the failures of the investment bank, Lehman Brothers, and the insurer, AIG, money market fund investors embarked on a flight-to-safety, through which they withdrew huge volumes of cash from prime funds and deposited huge volumes into government funds. Relatively suddenly, investors came to believe that financial entities might not be able to pay off their maturing commercial paper, which comprised a significant part of prime money market fund portfolios. And, in fact, the day after the bankruptcy of Lehman Brothers, the Reserve Primary Fund became the second fund in history to break the buck: the value of its sizable holdings of Lehman Brothers commercial paper had fallen so precipitously that the fund's NAV fell to 97 cents per share.[15] Fearing that investor flight from prime money market funds would exacerbate already stressed conditions in money markets, including the ability of financial entities to continue borrowing in CP markets, the Treasury instituted a program through which it guaranteed money market fund shares for one year, and the Federal Reserve created a facility that extended nonrecourse loans to banks collateralized by asset-backed commercial paper bought from money market funds.

[15]Reserve Primary Fund shareholders eventually received 99.1 cents per share.

After the crisis, the SEC changed several rules governing money market funds.[16] First, 2a-7 rules were tightened to increase the safety and improve the liquidity profiles of money market fund portfolios. Second, institutional prime and tax-exempt money market funds, as opposed to funds with only retail investors, must allow share price to *float* with the fund's NAV. Proponents of this change argue that floating NAVs raise awareness that fund values can fluctuate and may discourage withdrawals timed to precede a fund's breaking the buck. Opponents argue that money market fund investors, particularly institutional investors, are well aware of the risks; that floating NAVs have little to no bearing on flights-to-safety away from prime funds; and that floating NAVs significantly increase the accounting, operational, tax, and legal complexities of using money market funds for cash management. The third post-crisis change was that prime and tax-exempt money market funds had to have the power, under various stress conditions, to impose redemption fees of up to 2% and *gates* that prevent withdrawals for up to 10 business days in any 90-day period. Redemption fees, which are paid into the fund, are intended both to discourage investor withdrawals in a crisis and to recover losses from liquidating assets in a crisis to meet those withdrawals. Gates are intended to give funds a grace period in which to manage through stressed market conditions. If, however, a fund cannot restore stability by the end of its grace period, the fund is liquidated. Government funds, by the way, may choose to include the power to impose fees and gates, but very few have done so. While redemption fees and gates are intended to increase fund stability, they might actually encourage earlier, preemptive redemptions. In any case, both the floating NAV rule on institutional prime funds and the inclusion of redemption fees and gates on all prime funds went into effect in October 2016.

Figure O.12 shows the balances of money market funds over time, by sector. The timing of the steep drops in the balances of both prime and tax-exempt funds corresponds closely to the October 2016 effective date just mentioned. The inconvenience of floating NAVs and the potential loss of immediacy from redemption fees and gates clearly reduced the attractiveness of prime funds. In 2020, as part of a broad flight-to-quality brought on by the COVID pandemic and economic shutdowns, balances in government money market funds increased markedly. Over the whole of 2020, institutional prime fund balances stayed relatively constant at $600 billion, while retail prime fund balances fell gradually from about $500 billion to $200 billion. Prime balances did fall dramatically, however, in March 2020,

[16] As an aside, the Dodd-Frank Act of 2010 forbade the Treasury's use of its Exchange Stabilization Fund to guarantee money market funds. In March 2020, Congress lifted this ban through the end of 2020. Dodd-Frank also required that the Federal Reserve's emergency facilities for nonbanks have "broad-based eligibility" and be approved by the Secretary of the Treasury.

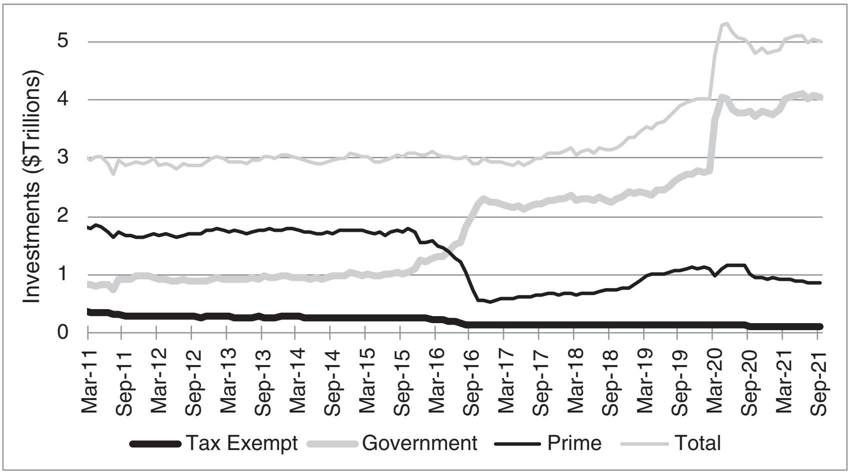

FIGURE 0.12 Balances in Money Market Funds, by Sector.
Source: US Money Market Fund Monitor, Office of Financial Research, US Department of the Treasury.

and institutional balances fell more than retail balances, but, due to swift action by the Treasury and Federal Reserve in support of financial markets in general and money market funds in particular, balances recovered relatively quickly. In any case, it seems that the possibility of redemption fees and gates in March 2020 did encourage preemptive withdrawals by prime investors and sales of assets by prime fund managers, who raised liquidity in order to avoid triggering fees and gates. Consequently, at the time of this writing, the SEC is revisiting fees and gates and considering other changes to the regulation of institutional prime funds.[17]

0.4 MONETARY POLICY WITH ABUNDANT RESERVES

To introduce the roles of the Federal Reserve system or "the Fed" as a modern central bank, Table O.4 shows its pre-crisis balance sheet, as of December 2007. One role, the creation and maintenance of a widely accepted national currency, is achieved by purchasing government bonds with that currency. The currency of the United States is comprised of green bills labeled as liabilities of the Federal Reserve, that is, as "Federal Reserve Notes." In terms of the balance sheet, therefore, outstanding currency is a liability and Treasury securities are assets. Put another way, currency is "backed" by government

[17]See Sidley (2022), "SEC Proposes New Rule Amendments for Money Market Funds," January 3.

TABLE 0.4 Balance Sheet of the Federal Reserve Banks, December 31, 2007, in $Billions.

Assets		Liabilities	
		Currency Outside Banks	773.9
Discount Window Loans	48.6	Bank Reserves and Vault Cash	75.5
Treasury Securities	740.6	Due to US Treasury	16.4
Repo	46.5	Repo	44.0
Other	115.2	Other	25.2
Total	950.9	Total	935.0

Sources: Board of Governors of the Federal Reserve System; and Author Calculations.

bonds. As shown in the table, $773.9 billion of currency outstanding was about 83% of Federal Reserve liabilities.

A second role of the central bank is to provide liquidity to banks under short-term distress. As mentioned earlier, bank assets, like loans, are relatively illiquid, while bank liabilities, like deposits or short-term borrowings from other financial institutions, may be due immediately. Therefore, a bank may find itself solvent but illiquid, that is, with assets of sufficient value to pay off its liabilities but without enough cash on hand to meet its immediate obligations. In these situations, a bank can borrow from the Fed through the *discount window* on any acceptable collateral. The discount window borrowings of a well-managed bank are expected to be infrequent and of relatively limited duration.

The third role discussed here, as set out in the Federal Reserve Act, is to "maintain long run growth of the monetary and credit aggregates . . . so as to promote effectively the goals of maximum employment, stable prices, and moderate long-term interest rates." Despite the three objectives, by the way, the Fed is often said to have a "dual mandate" of full employment and low inflation. In any case, two points are made before proceeding:

1. To ensure that banks have the resources to honor all requests to withdraw deposits that they are likely to receive, banks have been required to maintain cash and *reserves*, which are deposits held at a Federal Reserve bank. Until 2008, banks did not earn interest on their reserve balances.[18]
2. The Fed sets the total amount of reserves in the banking system. To add to reserves, the Fed might buy a Treasury bond from a bank. The bank's account at its Federal Reserve Bank is credited with the purchase

[18]The Federal Reserve System includes 12 Federal Reserve banks, spread around the country, and the Board of Governors, in Washington, D.C.

price – which increases bank reserves and the liabilities of the Fed – and the Treasury bond is added to the assets of the Fed. The Fed can also add to reserves by lending money to a bank through a repurchase agreement or repo.[19] The money is credited to the bank's account at the Fed, which is a reserve liability of the Fed, and the loan obligation is added to the Fed's assets. To reduce reserves, the Fed can do the opposite of these two transactions, that is, sell a bond to a bank or borrow money from a bank through a repo. In light of this discussion, the quantity of Treasury assets and net repo assets on the Fed's balance sheet enter into the determination of the total amount of reserves in the banking system.

Before the financial crisis of 2007–2009, reserves were scarce in the sense that banks traded reserves among themselves in the interbank *fed funds* market. Banks that needed reserves to satisfy their reserve requirements borrowed fed funds, while banks that had excess reserves, on which they did not earn interest, loaned fed funds at the market-determined fed funds rate. In this setting, if the Fed wanted to *ease* monetary conditions, to stimulate the economy (i.e., growth was too low and inflation not a threat), it would add reserves to the banking system, which – by increasing supply relative to demand in the market for reserves – would decrease the fed funds rate and likely increase the volume of bank loans to commercial enterprises. Similarly, if the Fed wanted to *tighten* monetary conditions, to slow the economy (e.g., growth was too inflationary), it would remove reserves from the system, which would increase the fed funds rate and likely decrease bank loans to commercial enterprises. Finally, the means by which the Fed changed reserves were called *open market operations* and consisted almost exclusively of lending and borrowing money through repurchase agreements.

In response to the financial crisis and ensuing Great Recession, the Fed tried to stimulate the economy as just described, by reducing the fed funds rate from 5.25% in September 2007 to a range of between 0% and 0.25% by December 2008. But having reduced interest rates to nearly zero and wanting to stimulate the economy even more forcefully, the Fed began what became known as *quantitative easing (QE)*: it purchased great quantities of US Treasury securities and "agency MBS," which are MBS issued by government agencies and GSEs. At the start of the program, purchases of MBS directly eased stresses arising from the overleveraged positions in mortgages and mortgage products that were at the heart of the crisis. Experts do not agree, however, on the mechanism by which the subsequent massive purchases of both Treasuries and MBS were to stimulate the economy, although

[19]As mentioned earlier, a repo is a loan collateralized by a security. In the context of the text, repos with the Fed are all collateralized by government-guaranteed securities.

possibilities include lowering intermediate- and long-term interest rates (rather than just the fed funds rate); pushing investors out of relatively safe Treasuries and agency MBS and into riskier products like corporate bonds and loans; and flooding banks with reserves. In fact, the chairman of the Fed at the time said a few years later that "the problem with QE is that it works in practice, but it doesn't work in theory."[20]

Figure O.13 shows the composition of the asset side of the Fed's balance sheet from June 2006 to June 2021. Note that the total of Fed assets is often referred to by market participants simply as the Fed's "balance sheet." In any case, during the financial crisis itself, from 2007 to 2009, assets were elevated by expanded repo lending, lending through the discount window, loans through emergency facilities that the Fed put in place at the time, and through swap lines, through which the Fed lends US dollars to foreign central banks, collateralized by foreign currency. While small in hindsight, this balance sheet expansion before QE was unprecedented, increasing from about \$900 billion at the start of 2007 to over \$2 trillion by the end of 2008. From then, the Fed fully engaged in QE by purchasing Treasuries and agency MBS, and the balance sheet grew dramatically. By the end of 2014, however, the Fed decided that economic conditions had improved enough to discontinue further stimulus. It stopped buying new assets, though it continued to roll over principal repayments from its holdings. From December 2015 to

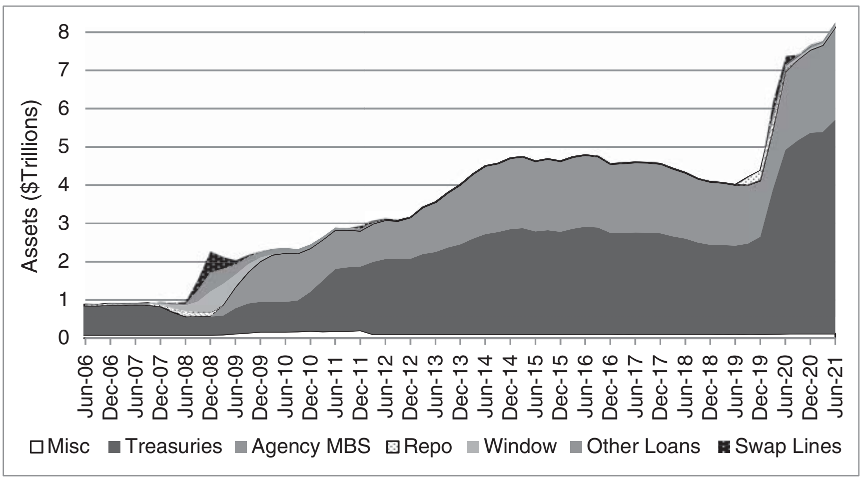

FIGURE O.13 Assets of Federal Reserve Banks, by Instrument.
Source: Board of Governors of the Federal Reserve System.

[20]Interview of Ben Bernanke by Liaquat Ahamed, "Central Banking after the Great Recession," Brookings, January 16, 2014.

summer 2019, it raised the target Fed funds rate from a range of 0% to 0.25% to a range of 2.25% to 2.50% and, in what was named the "normalization" of the balance sheet, began allowing assets to decline with principal repayments. The Fed was clear to point out, however, that the balance sheet would not decline to anywhere near its size before the financial crisis, because, as explained later, post-crisis monetary policy requires greater levels of reserves. Along these lines, in summer 2019, the Fed resumed the reinvestment of principal payments to maintain the size of the balance sheet. Then, in response to turmoil in the repo market in September 2019, further described later, the Fed started growing the balance sheet again. The Fed also judged that "implications of global developments for the economic outlook as well as muted inflation pressures" warranted lowering rates and reduced the Fed funds target range to 1.50% to 1.75% by November 2019. Finally, with the onset of the COVID pandemic and economic shutdowns, the Fed reduced rates back to the range of between 0% and 0.25% and aggressively bought Treasuries and MBS. As of June 2021, the Fed's balance sheet stood at $8.25 trillion, more than nine times its size before the financial crisis.

QE not only increased the size of the Fed's balance sheet but also changed its nature in a way that necessitated a new approach to implementing monetary policy. This is best described in terms of the liability side of the balance sheet, which is shown in Figure O.14 for three dates: the end of 2007, before the financial crisis; the end of 2019, before the pandemic; and, most recently, the end of the second quarter of 2021. At the end of 2007, consistent with the pre-crisis implementation of monetary policy, about 2% of the Fed's liabilities were bank reserves, and reserves were scarce. After years of QE, however, with Fed asset purchases adding to bank reserves, bank reserves rose to 37% of Fed liabilities by the end of 2019 and 43% by June 2021. Reserves had become abundant, and banks essentially stopped borrowing from and lending reserves to each other.[21] Consequently, in this new regime of abundant reserves, the Fed could no longer influence short-term rates by adding to or subtracting reserves from the banking system.

The Fed currently sets interest rates through two *administrative* rates, the *interest on reserve balances (IORB)* and the *reverse repo facility (RRP)* rate. The IORB is the rate that the Fed pays banks on their reserve deposits. Focusing on overnight, safe rates, that is, putting aside spreads due to longer terms, riskier investments, and fees, the IORB fixes the rate at which banks are willing to borrow and lend. No bank would borrow at a rate above the IORB: it would be less costly to use reserves and sacrifice earnings at the IORB. And no bank would lend at a rate below the IORB: it would be more profitable to keep reserves on deposit and earn the IORB.[22]

[21]Chapter 12 describes the current, limited trading in the fed funds market.

[22]The Fed recently substituted *Interest on Excess Reserves (IOER)* with IORB. Also, the system of setting rates described in the text is sometimes called a "floor" system,

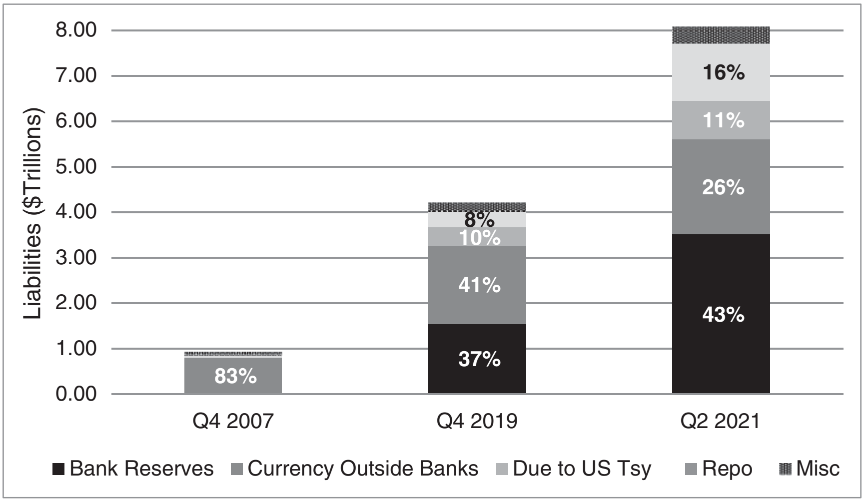

FIGURE 0.14 Liabilities of Federal Reserve Banks, by Instrument.
Source: Board of Governors of the Federal Reserve System.

If banks were the only participants in money markets, then setting the IORB would set overnight, safe rates in the system. But nonbanks with funds to lend might very well lend at rates below the IORB. First, only banks hold reserves and, therefore, only banks can lend to the Fed directly at the IORB. Second, because of post-crisis regulations, discussed later, banks are not willing to accept deposits from all comers and then hold those funds as their own reserves at the Fed. Therefore, other significant market participants, particularly money market funds, might very well lend funds at rates below the IORB. To prevent these lenders from pushing market rates below Fed targets, the Fed offers to pay a minimum rate to these participants through its RRP. More specifically, through the RRP, money market funds and some other entities can lend money to the Fed, taking Treasury securities as collateral, at the Fed's administered RRP rate.[23] It has turned out, in fact, that the facility has had to grow very large and very quickly to keep rates in the Fed's policy range. Figure O.14 shows that, as of June 2021, the Fed's repo liabilities of $1.3 trillion constituted 16% of the Fed's total liabilities.

---

because reserves are so plentiful that, between the relatively high discount window rate at which banks can borrow from the Fed, and the relatively low IORB at which banks can lend to the Fed, the market interest rate equals the lower IORB. In a "corridor" system, by contrast, reserves would be calibrated such that the market interest rate settles between the discount window rate and the IORB.

[23]In a reverse repo, a counterparty lends money and "reverses in" securities (see Chapter 10). Hence, the counterparties that are lending money to the Fed are doing reverse repo.

Putting these pieces together, the Fed sets rates in the regime of abundant reserves as follows: the Fed sets a target range for the market-determined fed funds rate; it sets the IORB to determine the rate at which banks are willing to borrow and lend; and it sets the RRP rate as a floor to the rate at which money market funds and others will lend. At the time of this writing, the fed funds target range is between 0% and 0.25%; the RRP rate is 0.05%; and the IORB is 0.15%. This means of implementing monetary policy has been successful in that the weighted-average of fed funds transactions has most recently been between 0.06% and 0.10%.

Policy implementation with abundant reserves has not been as successful, however, with respect to banks responding to market conditions without Fed assistance. Put another way, it has not been clear at what level reserves really are abundant. In mid-September 2019, reserves dropped by about 9% somewhat suddenly, from a variety of causes, including corporate tax payments and the settlement of new Treasury issues, both of which move funds from banks to the government. Banks might have been expected, with their abundant reserves, to supply funds to any market participant needing funds. But that did not happen. Instead, in the ensuing scramble for funds, when the Fed's target range was between 2% and 2.25%, the transaction-weighted repo rate on a particular day averaged 5.25%, and one trade was done at 10%. As mentioned earlier, the Fed responded swiftly by adding reserves, but the episode demonstrated that the determination of the quantity of reserves in a system of abundant reserves was not necessarily straightforward.

The reasons that reserves proved inadequate were traced to the interactions of monetary policy with bank regulation. First, since the financial crisis, banks have been subject to a leverage ratio, which requires that capital equal a minimum fraction of assets, regardless of the risk of those assets. In particular, reserves at the Fed, which are safer than any other instrument in the system, and repo lending on Treasury securities, which is extremely safe, are now both subject to capital requirements. Second, since the financial crisis, banks must hold enough *high-quality liquid assets (HQLA)* to meet certain scenarios of cash withdrawals and funding needs. Furthermore, while Treasury securities can be used to meet HQLA requirements, there appears to be a bank examiner preference for reserves. Third, a stigma remains for banks that have daylight overdrafts, that is, temporary negative balances at Fed accounts that have to be resolved by the end of the day. Returning to September 2019, then, banks were not willing to make additional repo loans, despite the relatively high rates available, because the loans would increase assets and capital requirements; use reserves; and risk daylight overdrafts.[24]

[24]For a fuller discussion of monetary policy with abundant reserves and repo markets in September 2019, see Copeland A., Duffie, D., and Yang, Y. (2021), "Reserves Were

A similar episode occurred in March 2020. News of the COVID pandemic and economic shutdowns resulted in disorder in Treasury markets, manifested through high costs of trading and high spreads between otherwise similar securities. These disturbances were much more severe than those in September 2019, but banks again seemed unwilling to use their resources, and the Fed again intervened by reducing the fed funds target range, offering unlimited repo loans, purchasing Treasuries and MBS, and temporarily excluding reserves, Treasuries, and Treasury repo from leverage ratio calculations.[25] As a more permanent response to the episodes of both September 2019 and March 2020, the Fed instituted a *standing repo facility (SRF)* in July 2021, through which primary dealers and bank counterparties can borrow money overnight from the Fed on government-backed collateral. To discourage the use of the facility except under stressed conditions, the borrowing rate is set above market rates. At the time of this writing, with the RRP rate at 0.05% and the IORB at 0.15%, the rate on the SRF is 0.25%.

Several policy issues arise with QE and the implementation of monetary policy in a regime of abundant reserves. First, the Fed was established and has traditionally accepted deposits only from banks. The idea was that the public deposits funds into banks, and banks decide to whom to make loans. In other words, the private banking sector was responsible for capital allocation decisions. The Fed has always held some quantity of government bonds, of course, because, in the current monetary system, as explained earlier, government bonds are the assets against which circulate the liability of central bank currency. But the purchase of much greater quantities of Treasuries – currency has fallen to only 26% of Fed liabilities – and the purchase of MBS by the Fed, are effectively capital allocation decisions. Furthermore, the acceptance by the Fed of large deposits from the public through the RRP has disintermediated banks as well. A second policy issue is that the existence of the RRP may encourage the public, in a crisis, to withdraw funds from banks and buy money market fund shares, to be deposited in the RRP. While withdrawing money from banks in cash has always been an option during crises, before the RRP this was very inconvenient for large sums. Third, to the extent that there is some limit on the size of a central bank's balance sheet, maintaining a large balance sheet in relatively normal times limits room for action in future times of stress. Fourth, the Fed is now paying banks interest on their reserves, which it was not doing before 2008.

Not So Ample After All," Staff Report Number 974, Federal Reserve Bank of New York, July; and Nelson, B. (2022), "The Fed Is Stuck on the Floor: Here's How It Can Get Up," Bank Policy Institute, January 11.

[25]With respect to this episode, see, for example, Baer, J. (2020), "The Day Coronavirus Nearly Broke the Financial Markets," *The Wall Street Journal*, May 20; and Duffie, D. (2020), "Still the World's Safe Haven?" Hutchins Center Working Paper #62, May.

This can be viewed as a cost to taxpayers when, for example, the Treasury is borrowing through three-month T-bills – and the Fed is investing some of its assets in those same T-bills – at five or six basis points, while the Fed is borrowing from banks through reserves at 15 basis points. And this issue could become more acute when the Fed eventually increases rates by increasing the IORB. In fact, one of the reasons that the Treasury has been keeping more of its cash balances or deposits at the Fed is to avoid the situation in which Treasury holds stores of cash as deposits at banks, earning a relatively low rate of interest, while the banks are depositing those fund at the Fed and earning the IORB.[26] Figure O.14 shows that Treasury deposits have grown to about 11% of Fed liabilities.

0.5 NEGATIVE RATES AND QE IN EUROPE AND JAPAN

Lending €1 at a negative rate means receiving less than €1 when the loan matures. Conversely, borrowing €1 at a negative rate means paying back less than €1 at maturity. Buying a bond at a negative yield means purchasing the bond for more than its face amount, receiving no interest over the life of the bond, and receiving only face amount at maturity.[27] Individuals with relatively small amounts of money can avoid lending at negative rates by keeping money in cash. For individuals and corporations with larger sums, however, holding cash is very cumbersome, and depositing funds at a bank at a modestly negative rate of interest may be the best available choice. There are also reasons to buy long-term bonds trading at a negative yield, despite their being guaranteed to lose money in nominal or euro terms. First, if bank deposit rates are negative, say at −0.50%, then purchasing a bond yielding −0.25% might be preferable to a deposit. Second, if the future is characterized by falling prices, that is, by deflation, then a negative yielding bond can offer a positive real return. For example, a bond yielding −1% when prices are falling at a rate of 2% is actually gaining 1% in real terms, that is, in purchasing power. Third, from a short-term trading

[26]And one of the reasons the Treasury has been holding greater cash balances is its wanting a cushion against the occasional political stalemates with respect to raising the debt ceiling, which, by limiting new borrowing, can leave the Treasury scrambling for cash.

[27]In theory, an investor could buy a negative yielding bond by paying its face amount, paying a periodic coupon, and then receiving the face amount at maturity. It is not feasible, however, for a government or corporate issuer to track down individual investors to collect coupon payments. Therefore, bonds with negative yields are sold as described in the text, with an initial price above par, a coupon of zero, and a return of par.

perspective, negative yielding bonds increase in price if yields fall. In other words, a trader makes money by buying a bond at a yield of −1% if market yields subsequently fall to −1.5%.

The Fed never lowered its target interest rate below zero as part of its easing program, but the European Central Bank (ECB) combined negative rates with QE starting in 2014, and the Bank of Japan did so in 2016. These policy decisions contributed to a peak of more than $18 trillion of global debt trading at negative yields in December 2020, and nearly as much as recently as summer 2021, with more than 50% of that volume in European bonds and about a third in Japanese bonds.[28] At the time of this writing, with central banks around the world expected to increase rates in response to inflation, the volume of negative yielding bonds is significantly lower, at less than $5 trillion in early 2022.

The Eurosystem refers to the ECB and the collection of national central banks in the euro area. Individual banks conduct transactions with and keep reserves at their respective national central banks, which, in turn, interact with the ECB. For simplicity, however, the discussion here is written as if banks trade directly with the ECB. The ECB targets interest rates by setting a *deposit facility rate*, which banks earn on their reserve deposits at the ECB, and a rate on *main refinancing operations*, which banks pay to borrow from the ECB through short-term repo transactions.[29] In its easing of monetary conditions, the ECB lowered these policy rates from 3.25% and 4.25%, respectively, in July 2008, to −0.50% and 0%, respectively, in September 2019. An individual bank might try to avoid earning a negative interest rate on its reserve deposits by lending out its reserves, which essentially passes them on to another bank. The banking system as a whole, however, cannot reduce the total amount of reserves supplied by the ECB and, therefore, cannot collectively avoid the negative rates on reserves. While the logic of reducing rates to stimulate economic activity can be extended to negative or even significantly negative rates, policymakers around the world have not been eager to go far along that path. Banks quickly passed negative rates on to their corporate depositors but have been reluctant or unable to do so for the vast majority of their retail depositors. Consequently, contrary to policy objectives, extended periods of negative rates seem to have reduced both bank profitability and lending activity.[30]

---

[28] See, for example, Flood, C. (2021), "In Charts: Bonds with Negative Yields Around the World," *Financial Times*, September 27.

[29] The ECB also sets a rate for borrowing through its *marginal lending facility*, which is the equivalent of the discount window at the Fed.

[30] See, for example, Beauregard, R., and Spiegel, M. (2020), "Commercial Banks Under Persistent Negative Rates," FRBSF Economic Letter 2020–29, Federal Reserve Bank of San Francisco, September 28; and Kowsmann, P. (2021), "Banks in Germany Tell Customers to Take Deposits Elsewhere," *The Wall Street Journal*, March 1.

To ease monetary conditions beyond reducing interest rates, the ECB turned both to loans to banks and to QE. Figure O.15 shows the assets of the ECB or, more precisely, the consolidated assets of the ECB and the national banks in the Eurosystem. Over the whole period, the multiplicative expansion of the balance sheet was similar to that at the Fed, shown in Figure O.13. The ECB began at a slower pace, however, and, at the start, placed a heavier reliance on bank loans. Before the financial crisis, the ECB loaned money to banks on a collateralized basis for a week through its *main refinancing operations (MROs)* and for three months through its *longer-term refinancing operations (LTROs)*. As the ECB wanted to ease monetary conditions, it offered longer-maturity LTROs, first up to a year and then up to three years. The logic was that banks have more flexibility to expand their lending if they are more certain of their source of funds. Then, starting in 2014, with the aim of making its loans even more stimulative, the ECB began *targeted long-term refinancing operations (TLTROs)*, which made four-year loans to banks in amounts based on the amounts that banks, in turn, loaned to their customers. At the time of this writing, MROs are for a week, LTROs for three months, and TLTROs for terms up to four years. As evident from Figure O.15, however, the ECB eventually expanded its balance sheet less through loans and more through QE, that is, through the purchase of securities. While government debt issues comprise the vast majority of these purchases, the ECB began purchasing nonbank corporate

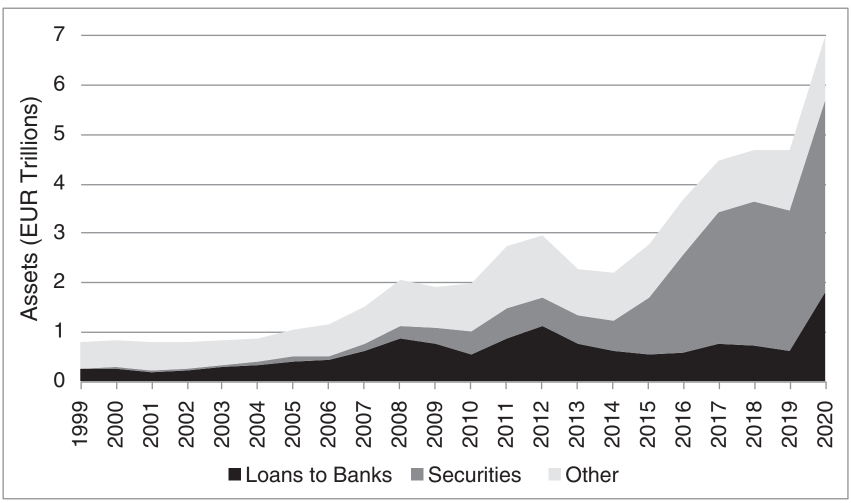

FIGURE O.15 Consolidated Balance Sheet of the Eurosystem, Assets.
Source: European Central Bank.

obligations in early 2016, and, as of early 2022, owns more than €350 billion of these securities.[31]

The ECB faces a unique challenge in implementing QE. European law prevents the ECB from funding individual European governments, or more broadly, from encouraging any unsound budget policies. In this spirit, to preserve ECB purchases as purely monetary rather than fiscal interventions, the ECB has aimed to purchase the bonds of various national governments in proportions to their *capital keys*, which reflect the sizes of their populations and economies. Furthermore, in 2015, the ECB limited itself to purchasing at most one third of the outstanding amount of any country's bonds. As QE purchases grew, however, keeping within these constraints has been difficult, and the ECB faced court challenges with respect to some of its decisions. One obstacle of growing significance has been that Germany has the greatest capital key, but a relatively small amount of debt outstanding. The ECB gave itself more leeway, therefore, for purchases under the Pandemic Emergency Purchase Programme in March 2020: the one-third limit would not be applied; the shortest eligible maturity would be 28 days rather than one year, which allowed for the purchase of short-term German government bills; bonds with less than investment-grade ratings would be eligible, which allowed for the purchase of Greek government debt; and some flexibility would be tolerated with respect to the capital keys constraint.[32]

Apart from allocating purchases across countries, the sheer magnitudes of ECB purchases have become challenging. For example, in June and July 2021, the ECB purchased €134.7 billion of government bonds from France, Germany, Italy, and Spain, while, over the same time period, the net issuance of those countries was only €89 billion. Furthermore, by the end of 2021, the ECB was expected to own more than 40% of all German and 40% of all Italian government bonds outstanding.[33]

One development that has made QE easier for the ECB is the relatively recent, large-scale issuance of debt by the European Commission (EC). EC

---

[31] These include €318 billion held under the Commercial Sector Purchase Programme (CSPP), as of early February 2022, and €50 billion under the Pandemic Emergency Purchase Programme (PEPP), as of the end of November 2021.

[32] "For purchases under the PEPP . . . the benchmark allocation . . . will be guided by the [capital] key[s] . . . A flexible approach to the composition of purchases . . . is nonetheless essential . . . " Source: Decision (EU) 2020/440 of the European Central Bank of 24 March 2020 on a temporary pandemic emergency purchase programme (ECB/2020/17), paragraph (5).

[33] Ainger, J. (2020), "One of the World's Top Bond Markets Slowly Capitulating to QE," December 9; Reuters (2021), "UPDATE-1-ECP Buys More Bonds Than Countries Sell to Cap Yields," August 2.

debt is an obligation of the European Union as a whole, and not the obligation of any particular country. While not unprecedented, such supranational European debt has never before been issued on the scale planned from 2020 to 2026. The EC is issuing €100 billion of SURE (Temporary Support to mitigate Unemployment Risks in an Emergency) bonds and €800 billion of NextGenerationEU bonds to finance expenditures aimed at supporting economic recovery from the COVID pandemic and economic shutdowns.[34] The total of these supranational issues is a modest, but not insignificant, fraction of the roughly €12.5 trillion of European government debt outstanding in 2021. From the perspective of the ECB, the debt issues of the EC not only provide more eligible bonds for purchase, but also bonds that are not subject to the cross-country purchase allocation constraints described earlier. EC debt purchases do face other constraints, however: at the time of this writing, the ECB can invest at most 10% of its portfolio in supranational debt and can hold no more than 50% of the debt of any one supranational issuer.

The Bank of Japan (BOJ) was the first central bank to introduce QE. Reacting to low growth and deflation, the BOJ had already, by 2001, pushed TONAR (Tokyo Overnight Average Rate), the Japanese interbank rate, to nearly zero. Looking for additional means of stimulus, the BOJ began its QE operations by purchasing Japanese government bonds (JGBs) with maturities greater than two years. Various additional lending and purchase programs were subsequently introduced, which included the purchase of corporate obligations, but, as shown in Figure O.16, its balance sheet did not increase at anything like the trajectory pursued by the Fed and ultimately the ECB until 2013, with the BOJ's *Quantitative and Qualitative Monetary Easing (QQE)*. This program committed not only to massive asset purchases, which were to include long-term JGBs, corporate bonds, exchange-traded funds, and equities, but also to lowering yields on long-term JGBs. QQE showed some promising economic results at first, but by 2016 the BOJ took further action. First, it "went negative," lowering the rate on some bank reserves to −0.10%. Second, it began *yield curve control* or *yield curve targeting*, in which it committed to buy (or sell) 10-year JGBs so as to keep their yields at approximately 0%. Yield curve control was effective in that the BOJ did not have to purchase so great a volume of bonds as previously to keep rates low. In other words, the mere threat of central bank trading to keep yields at 0% is enough to keep market-determined yields at that level. The effectiveness of this threat is manifested in Figure O.16 as a slower increase in BOJ assets from 2016 until the outbreak of the COVID pandemic

[34]The SURE bonds are "social bonds," or an environmental, social, and governance (ESG) debt instrument, and €250 billion of the NextGenerationEU bonds will be green bonds, used to fund eligible projects. This issue of green bonds will make the EC the world's largest issuer of green bonds.

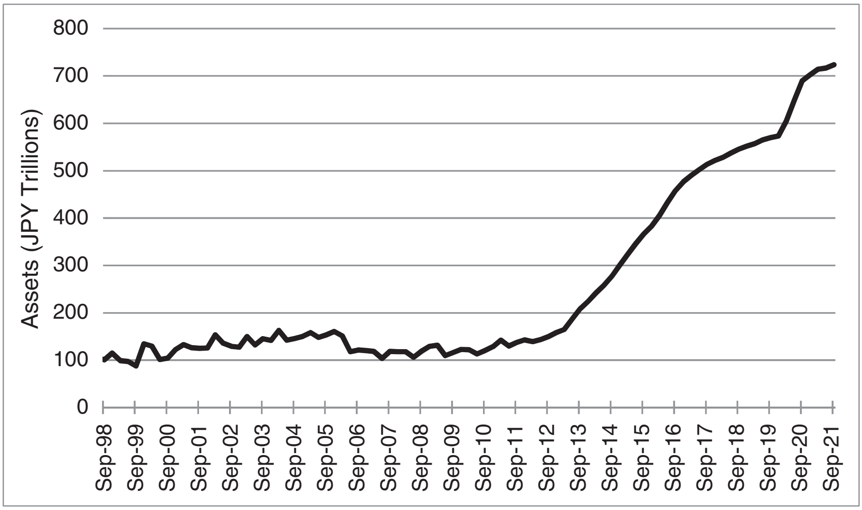

FIGURE 0.16 Assets of the Bank of Japan.
Source: Bank of Japan.

in 2020. While inflationary pressures in the United States and Europe, at the time of this writing, have pushed the Fed and ECB into more restrictive monetary postures, economic conditions and continued low inflation in Japan are keeping the BOJ committed to its expansionary policies.

0.6 TRADING AND LIQUIDITY

An important feature of any market is *liquidity*, loosely defined as the extent to which a market participant can trade a desired volume of securities at close to a prevailing market price. This section begins with some definitions of liquidity and some summary statistics across fixed income markets. Several trends impacting liquidity are discussed next, including electronification; dealer regulation and risk taking; the growth of *principal trading firms (PTFs)*; the growth of *exchange-traded funds (ETFs)*; and the growing concentration of asset managers. The section concludes with the implications of these trends with respect to liquidity and recent concerns about systemic fragility.

Because liquidity is hard to define precisely, there are many ways to measure it. The following are among the most common metrics:

- *Bid–ask spread* gives the difference between the higher ask price, at which a small amount of a security can be immediately purchased, and the lower bid price, at which a small amount of a security can be immediately sold. A market maker might expect to earn the bid–ask spread, on average, by buying securities from customers at the bid and selling to other customers at the ask. Equivalently, defining the *mid price* as the average of the bid and ask prices, and assuming the mid represents the fair market price, a customer can conceive of the cost of trading as half the bid–ask spread: buying at the ask instead of the mid and selling at the bid instead of the mid.

- *Volume* and *turnover* measure the amount of a security that trades over some time interval, as an absolute amount or as a percentage of the total outstanding amount of that security, respectively. A particular volume measure is *average daily volume (ADV)*, which averages realized daily volumes over some observation period. Volume and turnover both give insight into the size of a trade being contemplated. For example, it should not be too difficult to sell $1 million face amount of a bond when its ADV is $500 million. On the other hand, it will likely be quite challenging to sell 10% of an outstanding bond issue when its daily turnover is 2%.
- *Market depth* is a measurement of liquidity used in connection with a *central limit order book (CLOB)*. Through a CLOB, traders submit *limit orders* to buy a certain volume of a bond at a particular price (or less), or to sell a certain volume of a bond at a particular price (or more). Market depth is then defined in terms of the sum of the volumes of limit orders placed at prices from the mid to some level above or below the mid. For example, if the mid price is 100 and there are limit orders to buy $10 million face amount of bonds at prices of 99 or less, then a seller could immediately sell $10 million at 99. Unlike the bid–ask spread, therefore, market depth measures liquidity for orders of different sizes.
- *Price impact* or *implementation shortfall*. A buyer of a significant volume of a particular security can expect the market price to increase in the course of executing the purchases, as a seller of a significant volume can expect the market price to fall. Price impact measures the extent to which the market price can be expected to change as a result of a large transaction, and implementation shortfall measures the difference between the realized purchase or sale price on a large transaction relative to the market price before the transaction. Both price impact and implementation shortfall are typically estimated using data from past transactions. Asset managers and traders use the resulting models to estimate the transaction costs of their proposed trades. Dealers also use these models to quote prices at which they are willing to guarantee execution of large client orders.

The liquidity of fixed income securities varies dramatically across asset classes, across subclasses of those broader classes, and across individual securities within those subclasses. Using ADV, Table O.5 illustrates the variation of liquidity across broad asset classes. The ADVs in the table range

TABLE O.5 Average Daily Volumes (ADV) of Trading in Selected Markets, Fourth Quarter 2021, in $Billions.

Market	ADV
Treasuries	651.8
Agency MBS	251.2
Corporates	34.5
Munis	8.9
Agencies	2.4
ABS	1.3
Non-Agency MBS	1.0

Source: SIFMA (2022), "SIFMA Research Quarterly – 4Q21: US Fixed Income Markets – Issuance & Trading," January.

from over $650 billion for Treasuries and over $250 billion for agency MBS to about $9 billion for munis and even less for agency securities, asset-backed securities, and non-agency MBS. Variation of liquidity within these asset classes depends on the individual markets. In the Treasury market, for example, the most recently issued bonds in each maturity range are by far the most liquid (see, for example, Figure 11.1). In the corporate bond market, liquidity varies with age, credit quality, and issue size. Like Treasuries, the most recently issued corporates are the most liquid: one study reported that the monthly turnover of corporate bond issues declined from about 45% in the first month after issuance to about 13.5% in the second month and continues to decline thereafter. With respect to credit quality, investment-grade bonds are more liquid than less creditworthy high-yield bonds. In fact, 51.1% of the ADV of corporates in Table O.5 is due to trading in investment-grade bonds and only 17.5% to high-yield bonds. (The remaining ADV is from private placements, 25.8%, and convertibles, 5.6%.) Finally, with respect to issue size, one study showed that larger issues traded on a greater number of days in a year than smaller issues. To cite a few statistics, investment-grade bonds from issues greater than $1 billion trade, on average, on about 170 days a year. By contrast, investment-grade bonds from issues less than $250 million trade, on average, on fewer than 20 days in a year. Note, by the way, that describing liquidity in terms of the number of days on which a particular bond trades at all indicates the illiquidity of corporate bonds, at least relative to Treasuries. As a final example, the liquidity of muni issues also declines rapidly with age. Over the whole of 2019, the volume of munis that traded within a month of their issuance was about $585 billion. But the volume of munis that traded between one and three months of their issuance was only $138 billion, and between three and six months only $77 billion, etc. Furthermore, the total 2019 trading volume of all munis that had been issued 10 or more years previously was only about $91 billion.[35]

The text now turns to a discussion of trends affecting liquidity in fixed income markets.

Electronification

Equity and futures markets, which comprise relatively few distinct securities, migrated to electronic trading much more quickly than did bond markets, which comprise a vastly greater number of distinct issues. Traditionally, bonds were traded by *voice*. Customers traded with dealers, and dealers traded with each other, either directly or anonymously through interdealer brokers (IDBs). While a lot of bond trading is still by voice, electronic trading has been growing steadily. Dealer-to-customer (D2C) trades can take place on single-dealer or multi-dealer platforms, mostly through *request-for-quotes (RFQ)*. In this process, a customer requests dealers to provide price quotes to trade a particular size of a particular security or portfolio of securities. The customer then chooses with which dealer to trade.[36] A customer can also receive *direct streams* from dealers, through which dealers continuously send bid and offer prices for particular securities. The customer can then decide if and when to trade at the streamed prices. Dealer-to-dealer (D2D) trades can take place on interdealer electronic platforms, usually through a CLOB. Through a CLOB, traders submit limit orders to buy or sell a security, as described earlier, or they buy or sell that security immediately by lifting the offers or hitting the bids of other dealers who have submitted limit orders. PTFs have become more important in recent years and are discussed presently. PTFs often participate directly on interdealer platforms and also set up direct streams from individual dealers.

In the Treasury market, from February 2019 to May 2020, 65% of all trades were electronic. This percentage breaks up into 33% D2D trades, which nearly all trade electronically, and 67% DTC trades, of which 48% trade electronically. The electronic trades over this time period were 54% through CLOBs, 33% through RFQ, and 13% through streaming.

---

[35]Statistics cited in this paragraph are from: Blackrock (2017), "The Next Generation Bond Market;" Theisen, S. (2018), "Developments in Credit Market Liquidity," SEC FIMSAC Meeting, Citi, January 11; and MSRB (2019), "Municipal Securities Rulemaking Board 2019 Fact Book."

[36]Customers and dealers participating in RFQs have typically known each other's identities. Recently, however, anonymous RFQ platforms have gained traction.

In the corporate bond market, electronic trading of investment-grade bonds increased from 10% in 2011, to 19% in 2017, and 31% in 2020, and reached 38% in December 2020. Lower but growing percentages of high-yield bonds trade electronically, from 11% in 2017, to 21% in 2020, and to more than 25% in December 2020. Electronic trading of corporates is further ahead in Europe, growing from 39% of the market in 2017 to 47% in 2020.[37]

Dealer Regulation and Risk Taking

Dealers traditionally take risk in order to make markets. Customers can buy securities in significant size from dealers, because dealers are willing to bear the risks of buying and holding significant inventories in those securities. Customers can sell securities in significant size to dealers, because dealers are willing to bear the risks of holding and subsequently selling those securities. There are really two distinct risks here: market risk and execution risk. With respect to market risk, a dealer holding an inventory of corporate bonds can pretty easily hedge the risk of general interest rate changes, and even the risk of general changes to credit spreads, but hedging the unique credit risk of individual bonds can be difficult and expensive. With respect to execution risk, a dealer must typically "work out" of a large position. For example, after buying a large block of securities from a customer, the dealer cannot usually sell that whole block at once: the loss from the resulting price impact would likely wipe out any spread or fee that the dealer had charged the customer. Therefore, the dealer usually sells the larger position gradually, in smaller pieces, running the risk that other market participants happen to jump in to sell the same security at the same time, and also extending any idiosyncratic market risks of the position.

Dealers in banking entities were subjected to more stringent regulation after the financial crisis of 2007–2009. In particular, capital requirements were increased, both adjusted for risk and as a raw percentage of assets; liquidity requirements were formally added; and trading purely for the account of the bank, rather than for facilitating customer transactions, was – with the exception of Treasury bond trading – severely restricted by the Volcker rule.

[37]Mackenzie, M. (2021), "Pandemic's Digital Push Shows Future of Bond Trading," *Financial Times*, November 7; McPartland, K. (2021), "All-to-All Trading Takes Hold in Corporate Bonds," Greenwich Associates, Second Quarter; McPartland, K., Monahan, K., and Swanson, S. (2020), "US Capital Markets Performance During COVID," Greenwich Associates, December 22; Wiltermuth, J. (2021), "Electronic Trading in US Corporate Bonds Is Finally Taking Off," MarketWatch, July 14; and Author Calculations.

Because of this new regulatory regime, and perhaps also because of their own choices to reduce risk, banking entities are less willing to hold securities inventory and less willing to bear the risks of making markets. Figures O.17 and O.18 illustrate this point in the US Treasury and corporate bond markets, respectively.

Figure O.17 shows that the amount of 10-year Treasury notes sold at auction has been increasing steadily, and particularly rapidly in recent years, while, at the same time, dealers – who are mostly within banking entities – have been purchasing lower and lower percentages of those sales. Both the issue sizes and allotments to dealers, given in the figures by the very light gray lines, vary significantly from month to month, mostly because of the auction cycle of new and reopening sales described previously. To show the trend more clearly, therefore, the solid lines are centered seven-month averages of the underlying series. In any case, in the last few month of the observation window, the allotment to dealers is well under 15%. Figure O.18 shows that primary dealers are holding much smaller net positions in corporate bonds than before the crisis, while the total amount of corporate bonds outstanding has increased dramatically.[38]

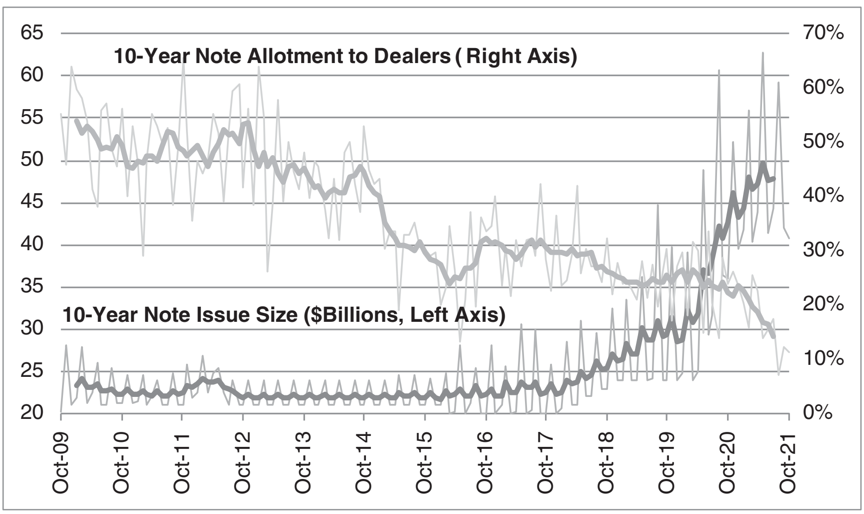

FIGURE O.17 10-Year Treasury Note Auction Size and Allotment to Dealers.
Sources: Investor Class Auction Allotments, US Department of the Treasury; and Author Calculations.

[38]Primary dealers commit to participate in Treasury auctions and are trading counterparties and market makers for the Fed.

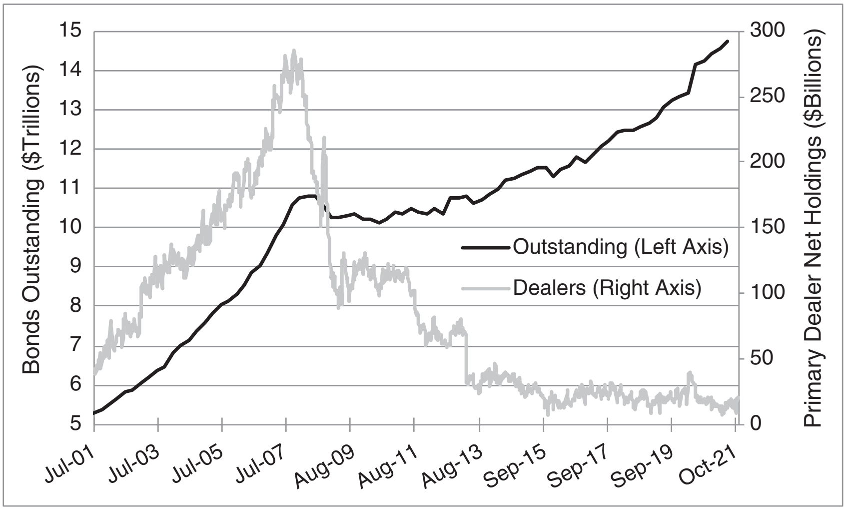

FIGURE 0.18 Primary Dealer Net Holdings of Corporate Bonds and Corporate Bonds Outstanding.
Sources: Primary Dealer Statistics, Federal Reserve Bank of New York; and Financial Accounts of the United States, Board of Governors of the Federal Reserve.

Growth of PTFs

In the Treasury market, the reduction of dealer market making has been made up, at least in part, by PTFs. These firms typically trade to make money on their accounts and try to make relatively small amounts of money on each of very many trades. They are often high-frequency trading firms (HFTs) as well, but do not have to be. In any case, the PTF business model or trading style relies heavily on electronification, and the advance of electronification is in good part due to demand from PTFs. Market participants were shocked in 2015 by an inadvertent leak of the top 10 firms by trading volume on the largest platform for the electronic trading of Treasuries: only two were banks. Figure O.19 uses more recent data to show how PTFs have grown in prominence in the Treasury market. Over the sample period, PTFs and dealers each comprised 48% of trading in the interdealer broker market (with buy-side firms, like asset managers, comprising the remainder). Furthermore, true to their usual nature, about 93% of PTF trades were electronic, compared with 58% of dealer trades. Note, by contrast, that only 18% of buy-side trades were electronic.

Growth of ETFs

Shares of an ETF represent fractional holdings of the fund's underlying portfolio and trade on an exchange. *Authorized Participants (APs)* of an ETF can

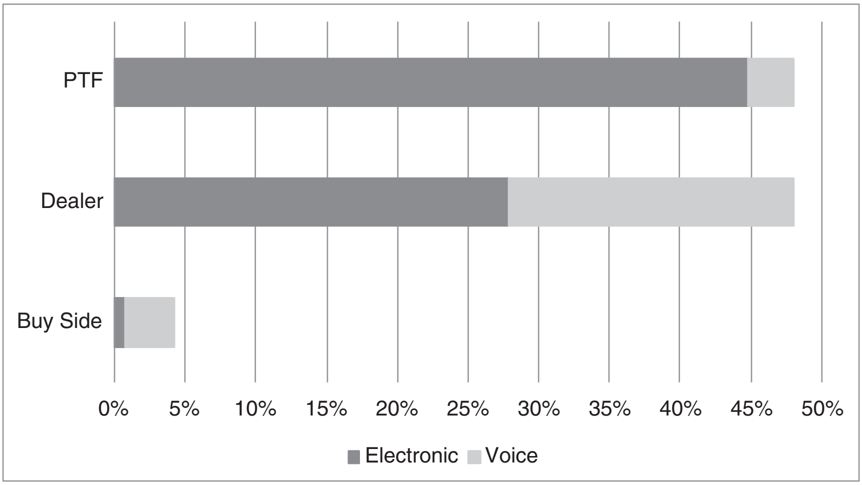

FIGURE 0.19 Percentages of Trading Volume in the Interdealer Broker Markets for Treasuries, by Participant Type and Trading Protocol, April to December 2019.
Source: Harkrader, J., and Puglia, M. (2020), "Principal Trading Firm Activity in Treasury Cash Markets," *FEDS Notes*, Board of Governors of the Federal Reserve System, August 4.

create ETF shares by buying underlying assets, adding them to the fund's portfolio, and selling ETF shares against those additional assets. APs can also reduce the number of outstanding ETF shares by taking assets from the underlying portfolio, selling them, and purchasing existing ETF shares corresponding to the sold assets. An ETF has a NAV, usually expressed as a value per share, which is an estimate of the value of the underlying portfolio. However, investors buy or sell shares of an ETF on an exchange at a market-determined price, which may be greater than, equal to, or less than the NAV. Unlike APs, investors cannot exchange their shares for securities in the underlying portfolio, nor can they deliver securities in exchange for new ETF shares.

ETFs in general, including fixed income ETFs, have experienced phenomenal growth. The first fixed income ETF was listed in 2002, and, by the end of 2020, assets under management in global bond ETFs exceeded $1.5 trillion.[39] An important motivation for bond ETFs is the creation of liquid trading vehicles from relatively illiquid assets. As discussed earlier, in part because there are so many distinct corporate bond issues, individual corporate bonds are relatively illiquid, particularly as they age. By contrast,

[39]iShares by Blackrock (2021), "By the Numbers: New Data Behind the Bond ETF Primary Process."

a small number of large and diversified portfolios, traded on exchanges and traded electronically in the form of ETFs, can be very liquid. Furthermore, growing liquidity and electronic trading in ETFs drive more electronic trading of corporate bonds, which further improves ETF liquidity, and so forth. In fact, a market has developed for trading portfolios comprising bonds that are all or mostly all held by ETFs.[40] In any case, Figure O.20 illustrates the liquidity advantages of ETFs. For each asset class shown, ETF assets under management (AUM) are a small fraction of total amounts outstanding. At the same time, however, ETF ADV is a large fraction of total ADV. For example, ETFs hold only 4.8% of high-yield corporate bonds outstanding, but the ADVs of these ETFs comprise 24.3% of the ADV of all such bonds. In other words, high-yield corporate bond ETFs are significantly more liquid than high-yield corporate bonds.

Skeptics have argued that day-to-day ETF liquidity lulls investors into a false sense of security, in the sense that, under stressed market conditions, ETF liquidity will revert to the liquidity of the underlying bonds and essentially disappear. This theory was tested in March 2020, when markets roiled from the COVID pandemic and economic shutdowns. Underlying bond market liquidity did evaporate, with even primary issues of both

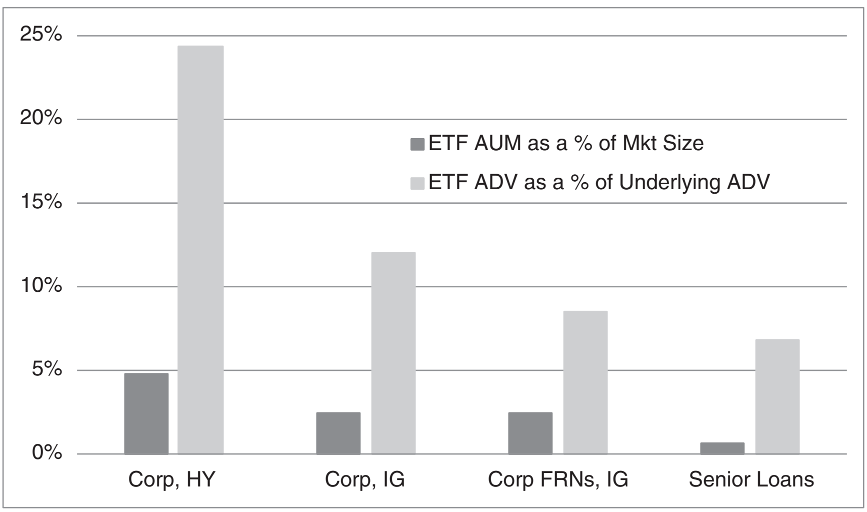

FIGURE O.20 Exchange-Traded Fund AUM and ADV as a Percentage of Market AUM and ADV.
Source: State Street Global Advisors (2020), "Fixed Income ETFs: Fact versus Fiction (Answers on Fund Structure, Liquidity, Trading and Performance)," June.

[40] Mackenzie, M. (2021), "Pandemic's Digital Push Shows Future of Bond Trading," *Financial Times*, November 7.

investment-grade and high-yield bonds coming to a halt. Over the same time, however, ETF trading volumes increased dramatically. In other words, investors could not trade individual bonds, but could trade ETFs. The pricing of ETFs over the period, however, surprised some observers. As bond prices plummeted, ETFs traded at significant discounts from NAV; that is, the prices of the ETFs were very much below the estimated value of their underlying portfolios. Similarly, when bond prices rapidly recovered, ETFs traded at significant premiums to NAV, that is, at prices well above estimated portfolio values. Critics took this pricing behavior as evidence that ETFs were not working well, in the sense that investors could not trade shares at their true values. But by the definition of illiquidity during stress events, "true values" of individual bonds cannot be accurately determined at those times. It is most likely, therefore, that observed discounts from NAV resulted from stale bond prices: as prices swiftly fell with nearly no liquidity, bond prices recorded for the purposes of estimating NAV were stale, higher prices. Similarly, as prices rebounded quickly with little liquidity, NAV was estimated with stale, lower prices, making it seem that ETFs traded at a premium. Supporting this narrative was the fact that when prices were falling and ETFs were selling at a discount to NAV, investors were nevertheless giving bonds to APs to create new ETFs. Despite the apparent loss in value, investors could sell their bonds most efficiently by first exchanging them into ETFs.[41]

Growing Concentration of Asset Managers

Concentration in the asset management industry has been growing for some time. Assets under management at the top 20 global managers were 29% of total assets in 1995, 38% in 2000, and 43% in 2020. Furthermore, in 2020, the top 10 global managers managed 31% of total assets, and the top five, 21%. With respect to flows, a 2018 study found that 42% of asset management trading in investment-grade corporate bonds came from the top five managers.[42] Increasing concentration in asset management implies a growing demand in the market for larger trades or, equivalently,

[41]For accounts of ETFs during March 2020, see, for example: Aramonte, S. and Avalos, F. (2020), "The recent distress in corporate bond markets: cues from ETFs," *BIS Bulletin* No. 6, April 14; Levine, M. (2020), "Money Stuff: The Bull Market Caught a Virus," Liquidity Illusion Illusion, *Bloomberg*, March 12; and S&P Global Ratings (2020), "Credit Trends: How ETFs Contributed to Liquidity and Price Discovery in the Recent Market Dislocation," July 8.

[42]McPartland, K. (2019), "The Challenge of Trading Corporate Bonds Electronically," Greenwich Associates, Q2; Thinking Ahead Institute (2020), "Global Asset Manager AuM tops US$100 Trillion for the First Time," Willis Towers Watson, October 19; and Author Calculations.

a growing amount of execution risk from having to break up large orders into smaller transactions.

Summary of Trends

The trends just presented can be combined into the following narrative. Dealers are allocating less risk to market making, while asset managers need to execute larger and larger trades. Therefore, dealers focus on their largest and best clients, pushing others to electronic platforms. The growth of PTFs makes up for some of the liquidity withdrawn by dealers, but by providing frequent liquidity on small trades rather than by assuming the execution risk of larger orders. The net results of these intersecting trends, therefore, are as follows:

- Liquidity has become plentiful for small trades in relatively liquid, electronic markets.
- Liquidity can still be challenging for larger trades, even in relatively liquid markets, and, of course, for trades in less liquid markets.
- More customers must bear the execution risk of their own trades, instead of paying dealers to assume this risk.

Liquidity and Fragility

The changing nature of liquidity in bond markets has raised the issue of market resilience or, its reverse, market fragility. In particular, some believe that liquidity provision by PTFs is less stable than by dealers. The argument is that PTFs will shut down in stress conditions, so as to avoid any losses on their own accounts, while dealers will continue to make markets for their clients, with whom they have ongoing and valuable business relationships. The "flash rally" in the Treasury market on October 15, 2014, provided a case study in which to examine this argument.

On that day, there was a release of retail sales data at 8:30am, but it is generally agreed that the news was not particularly surprising. Nevertheless, Treasury yields dropped relatively steeply over the next hour or so. Then, over the 12 minutes from 9:33am to 9:45am, the yield on the 10-year Treasury fell by 16 basis points and then rose by 16 basis points (i.e., prices rose by a lot and then fell by a lot). This was an extremely large move for such a short time period – the daily standard deviation of the 10-year yield was perhaps four or five basis points *per day*. Was any particular group of market participants to blame for this flash rally? Does the event forewarn of worse occurrences to come?

Subsequent analysis showed that over the 12 minutes in question, trading volume increased dramatically and market depth dropped precipitously. In other words, trading continued throughout the interval, and in large quantities, but through very many small orders and a consistent replenishment of limit orders. Market depth provided by both PTFs and dealers fell during the window. While market depth supplied by PTFs fell by a much larger percentage, PTFs supplied much more of the total throughout. Furthermore, PTFs did not change their bid–ask spreads by much, while dealers did at times. A joint study of regulators concluded that: "In very broad terms . . . PTFs, as a group, reacted . . . primarily by reducing limit order quantities, while the bank-dealers reacted by widening bid–ask spreads and, for brief periods of time, removing their offers to sell securities."[43] Evidence from the flash rally, therefore, is consistent with the different nature of liquidity provision by PTFs and dealers; is not consistent with PTFs closing shop during a stress event; and is consistent with a liquidity regime that is shifting execution risk to the buy side.

There was a similar flash event on February 25, 2021, though this time Treasury prices fell significantly and then recovered in a short amount of time. The event was again characterized by sharply reduced market depth and sharply elevated trading volumes, that is, by rapid replenishment of limit orders.[44] The extent to which markets are susceptible to similar and perhaps worse occurrences remains a topic of concern.

[43] Joint Staff Report (2015), "The US Treasury Market on October 15, 2014," p. 28.
[44] Aronovich, A., Dobrev, D., and Meldrum, A. (2021), "The Treasury Market Flash Event of February 25, 2021," *FEDS Notes*, Board of Governors of the Federal Reserve System, May 14.

CHAPTER 1

Prices, Discount Factors, and Arbitrage

This chapter begins by introducing the cash flows of fixed-rate, government coupon bonds. It shows that prices of these bonds can be used to extract *discount factors*, which are the market prices of one unit of currency to be received on various dates in the future.

Relying on a principle known as *the law of one price*, discount factors extracted from a particular group of bonds can be used to price bonds that are not part of that original group. Furthermore, a particularly persuasive relative pricing methodology, known as *arbitrage pricing*, turns out to be mathematically identical to pricing with discount factors. Hence, discounting can rightly be used and regarded as shorthand for arbitrage pricing.

Market prices on a single, fixed date are used to illustrate that the law of one price and arbitrage pricing describe the US Treasury market relatively well, but not perfectly. Bonds are not commodities: their prices reflect supply and demand characteristics that are not fully captured by their scheduled cash flows. The US Treasury's *Separate Trading of Registered Interest and Principal of Securities (STRIPS)* are introduced next, both as a topic of independent interest and as an additional illustration of the complex realities of pricing in bond markets.

The chapter concludes with accrued interest and day-count conventions, which are used throughout fixed income markets and throughout this book.

For clarity of exposition, prices and examples in this chapter are all as of the close of business on Friday, May 14, 2021. Also, because transactions in the US Treasury market typically *settle* "T+1," that is, bonds and cash are exchanged one business day after the trade date, all trades in the examples settle on Monday, May 17, 2021.

1.1 GOVERNMENT COUPON BONDS

The cash flows from government coupon bonds are defined by *coupon rate*; *maturity date*; and, synonymously, *face amount*, *principal amount*, or *par value*. For example, in May 2014 the US Treasury sold a bond with a coupon rate of 2.5% and a maturity date of May 15, 2024. Purchasing $1 million face amount of these "2.5s of 05/15/2024" entitles the buyer, as of settlement in mid-May 2021, to the schedule of payments depicted in Table 1.1. More specifically, the Treasury promises to make a coupon payment every six months equal to half the bond's annual coupon rate of 2.5% times the face amount, that is, $2.5\%/2 \times \$1,000,000 = \$12,500$. Finally, on the maturity date of May 15, 2024, in addition to the coupon payment on that date, the Treasury promises to pay the bond's face amount of $1,000,000.[1]

This chapter restricts attention to US Treasury bonds, but the analytics presented here apply easily to bonds issued by other countries, because cash flows across sovereign bonds differ mostly with respect to the frequency of coupon payments. Government bonds in France and Germany, for example, make annual coupon payments, while those in Italy, Japan, and the United Kingdom pay semiannually, as in the United States.

Returning to the US Treasury market, Table 1.2 reports the prices of selected US Treasury bonds as of May 14, 2021. Per market convention, bond prices are always quoted per 100 face amount. Also, each price in the table is a *full* or *invoice* ask price, that is, the total price at which traders stand ready to sell that particular bond. (The division of full price into a *flat* or *quoted* price and *accrued interest* will be explained later in the chapter.) From the third row of the table, then, for example, purchasing the 1.625s

TABLE 1.1 Cash Flows of $1 Million Face Amount of the 2.5s of 05/15/2024. Entries Are in Dollars.

Date	Coupon Payment	Principal Payment
11/15/2021	12,500	
05/15/2022	12,500	
11/15/2022	12,500	
05/15/2023	12,500	
11/15/2023	12,500	
05/15/2024	12,500	1,000,000

[1]Scheduled bond payments that do not fall on a business day are made the following business day. For example, because May 15, 2022, is a Sunday, the coupon payment due that day would actually be paid on Monday, May 16, 2022.

TABLE 1.2 Prices of Selected US Treasury Bonds Maturing on November 15 or May 15 of a Given Year, as of May 14, 2021. Coupons Are in Percent.

Coupon	Maturity	Price
2.875	11/15/2021	101.4297
2.125	05/15/2022	102.0662
1.625	11/15/2022	102.2862
0.125	05/15/2023	99.9538
0.250	11/15/2023	100.0795
0.250	05/15/2024	99.7670
2.250	11/15/2024	106.3091

of 11/15/2022 costs 102.2862 per 100 face amount or \$102,286,200 for \$100,000,000 face amount.

The bonds in Table 1.2 are selected from the broader list of US Treasuries for two reasons. First, it is convenient for the computations in the next section that bonds mature in approximate six-month intervals. Second, when two or more Treasury bonds mature on a given date, the most recently issued bond is chosen because it is likely to be more liquid and, consequently, its quoted price more reliable. For example, the 0.25s of 05/15/2024, which were issued in 2021, are chosen over the 2.5s of 05/15/2024, which were featured in Table 1.1 and issued in 2014.

1.2 DISCOUNT FACTORS

The *discount factor* for a particular term gives the value today, or the *present value*, of one unit of currency to be received at the end of that term. Denote the discount factor for t years by $d(t)$. Then, for example, if $d(3.0) = 0.99$, the present value of \$1 to be received in three years is 99 cents.

Because Treasury bonds promise future cash flows, discount factors can be extracted from Treasury bond prices. In fact, the rows of Table 1.2 can be used to write equations that relate bond prices to discount factors. Starting with the first row of the table, that is, with the bond maturing in six months,

$$101.4297 = \left(100 + \frac{2.875}{2}\right) d(0.5) \tag{1.1}$$

In words, the price of the 2.875s of 11/15/2021 equals the present value of its future cash flows, or, more precisely, the sum of its principal and coupon cash flows multiplied by the discount factor for funds to be received in six months. Solving, $d(.5) = 0.999923$.[2]

By the same reasoning, the equations relating prices to discount factors can be written for the other bonds in Table 1.2. For the next two bonds, which mature in one year and in 1.5 years, respectively,

$$102.0662 = \frac{2.125}{2}d(0.5) + \left(100 + \frac{2.125}{2}\right)d(1.0) \tag{1.2}$$

$$102.2862 = \frac{1.625}{2}d(0.5) + \frac{1.625}{2}d(1.0) + \left(100 + \frac{1.625}{2}\right)d(1.5) \tag{1.3}$$

Given the solution for $d(0.5)$, from Equation (1.1), Equation (1.2) can be solved for $d(1)$. Then, given the solutions for both $d(0.5)$ and $d(1.0)$, Equation (1.3) can be solved for $d(1.5)$. Continuing in this way through the rows of Table 1.2 solves for the discount factors, in six-month intervals, out to three and one-half years. Table 1.3 reports these results. That these discount factors fall as term increases reflects the time value of money: the longer a payment of $1 is delayed, the less it is worth today.[3]

Figure 1.1 shows discount factors, in six-month intervals, from the bonds listed in Table 1.2 and other bonds, not shown here, which have maturity dates on November 15 and May 15 out to 10 years. Discount factors can again be seen to decrease with term, with the value of $1 to be received in 10 years worth less than 85 cents as of the pricing date.

TABLE 1.3 Discount Factors Derived from Bonds Listed in Table 1.2.

Term	Discount Factor
0.5	0.999923
1.0	0.999419
1.5	0.998504
2.0	0.997041
2.5	0.994558
3.0	0.990195
3.5	0.984742

[2]For ease of exposition, this section assumes that the coupon and maturity dates of the bonds listed in Table 1.2 are exactly 0.5 years apart. A more precise construction of discount factors would use the exact number of days between cash flow dates.
[3]Discount factors will not necessarily fall with term if interest rates are negative.

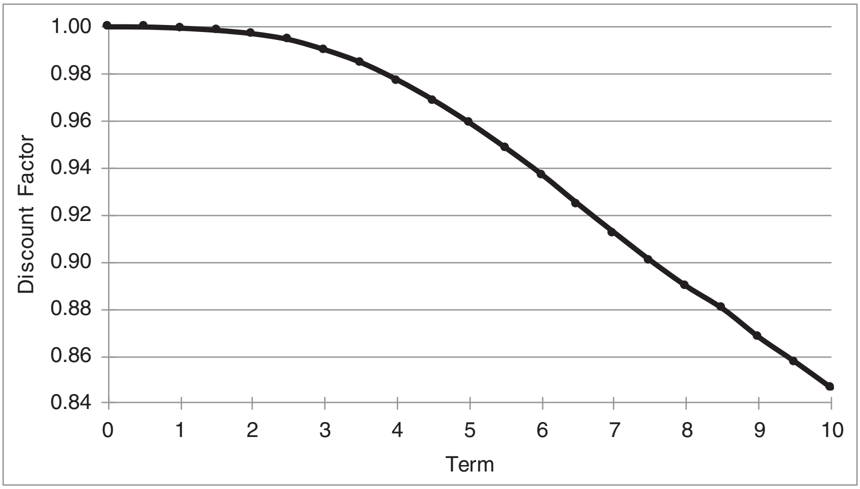

FIGURE 1.1 Discount Factors Derived from Selected US Treasury Bonds, as of May 14, 2021.

1.3 THE LAW OF ONE PRICE

The US Treasury 1.75s of 05/15/2022 mature in mid-May, approximately one year from the settlement date of the examples of this chapter but are not included among the bonds in Table 1.2. How should the 1.75s of 05/15/2022 be priced? A natural answer is to apply the discount factors of Table 1.3 even though the 1.75s of 05/15/2022 are not used to construct those discount factors. After all, because all of these bonds are obligations of the US Treasury, it seems reasonable to assume as a first approximation that the present value of receiving $1 from the Treasury on some future date does not depend on which particular bond pays that $1. This reasoning is known as the *law of one price*: absent confounding factors (e.g., liquidity, financing, taxes, credit risk), identical cash flows should sell for the same price.

According to the law of one price, then, the price of the 1.75s of 05/15/2022 should be,

$$\frac{1.75}{2} \times 0.999923 + \left(100 + \frac{1.75}{2}\right) \times 0.999419 = 101.691 \qquad (1.4)$$

where the bond's present value is calculated as the sum of each cash flow multiplied by its corresponding discount factor from Table 1.3. As it turns out, the market price of this bond is 101.693, which is nearly equal to the prediction of the law of one price in Equation (1.4).

Table 1.4 compares market prices with present values for all US Treasury bonds that mature on November 15 or May 15 out to November 15, 2024.

TABLE 1.4 Treasury Bond Prices Versus Present Values Using the Discount Factors of Table 1.3. Bonds in Bold Are Used to Derive those Discount Factors. Coupons Are in Percent.

Term	Coupon	Maturity	Market Price	Present Value	Rich(+)/ Cheap(−)	Issue Date
0.5	**2.875**	**11/15/2021**	**101.4297**	**101.4297**	**0.0000**	**11/15/2018**
0.5	2.000	11/15/2021	100.9952	100.9922	0.0030	11/15/2011
0.5	8.000	11/15/2021	104.0904	103.9920	0.0984	11/15/1991
1.0	**2.125**	**05/15/2022**	**102.0662**	**102.0662**	**0.0000**	**05/15/2019**
1.0	1.750	05/15/2022	101.6931	101.6914	0.0017	05/15/2012
1.5	**1.625**	**11/15/2022**	**102.2862**	**102.2862**	**0.0000**	**11/15/2012**
1.5	7.625	11/15/2022	111.3969	111.2797	0.1172	11/16/1992
2.0	**0.125**	**05/15/2023**	**99.9538**	**99.9538**	**0.0000**	**05/15/2020**
2.0	1.750	5/15/2023	103.1970	103.1997	−0.0026	05/15/2013
2.5	**0.250**	**11/15/2023**	**100.0795**	**100.0795**	**0.0000**	**11/16/2020**
2.5	2.750	11/15/2023	106.3040	106.3163	−0.0123	11/15/2013
3.0	**0.250**	**05/15/2024**	**99.7670**	**99.7670**	**0.0000**	**05/17/2021**
3.0	2.500	05/15/2024	106.5448	106.4941	0.0508	05/15/2014
3.5	**2.250**	**11/15/2024**	**106.3091**	**106.3091**	**0.0000**	**11/17/2014**
3.5	7.500	11/15/2024	124.8220	124.5906	0.2314	08/15/1994

The "Present Value" column is computed along the lines of Equation (1.4), giving the prices of the bonds as predicted by the law of one price. The column "Rich (+) / Cheap (−)" is the difference between the market price and the present value. Bonds with prices greater than that predicted by a model, like the law of one price, are said to be trading *rich*, while bonds with prices less than predicted by the model are said to be trading *cheap*.

The bonds in bold in Table 1.4 are those listed in Table 1.2 and used to compute the discount factors in Table 1.3. Therefore, these bonds are *fair*, that is, neither rich nor cheap, relative to those discount factors. The prices of the other bonds in Table 1.4, however, can be either rich or cheap relative to those discount factors or, equivalently, relative to the prices of the bonds in bold.

The table, as a whole, more or less supports the law of one price. Most of the deviations of present values from market prices are very small, and even the largest deviation, the 23-cent richness of the 7.5s of 11/15/2024, is less than 0.2% of the bond's market price.

Across the bonds in Table 1.4, those with the largest deviations are the older, high-coupon bonds, for example, the 8s of 11/15/2021, issued in 1991; and the 7.625s of 11/15/2022, issued in 1992. Furthermore, these bonds all trade rich relative to the bonds in bold. This is somewhat surprising because, historically, older bonds, which tend to be relatively illiquid, have tended to trade cheap relative to other bonds. Perhaps in the current, low-rate environment, some investors have a strong preference for income and are willing to pay more than fair value for bonds paying high coupons.

While the law of one price is generally supported by Table 1.4, it is natural to ask whether a trader or investor could profit by buying cheap bonds and simultaneously selling fairly priced bonds; by selling rich bonds and simultaneously buying fairly prices bonds; or – so as to profit on both sides of the trade – by buying cheap bonds and simultaneously selling rich bonds. The next section addresses this question.

1.4 ARBITRAGE AND THE LAW OF ONE PRICE

While the law of one price is intuitive, its real justification rests on a stronger foundation. It turns out that a deviation from the law of one price implies the existence of an *arbitrage opportunity*, that is, a trade that generates profit without any chance of losing money.[4] But because arbitrageurs would flock toward any such trade, market prices can be expected to adjust quickly so as to rule out any significant deviations from the law of one price. Put another way, arbitrage activity can be expected to enforce the law of one price.

To make this argument more concrete, consider an arbitrage trade based on the richness of the 7.625s of 11/15/2022, as reported in Table 1.4. More specifically, sell or *short*[5] the 7.625s of 11/15/2022 and simultaneously buy a portfolio of bonds, from among the bonds in bold in Table 1.4, which perfectly replicates the cash flows of the 7.625s of 11/15/2022. Because the 7.625s of 11/15/2022 are rich relative to the bonds in bold, selling the former and buying a portfolio of the latter, in a way that generates no future cash flows, should generate a riskless profit. Table 1.5 describes this *replicating portfolio* and this arbitrage trade in more detail.

Columns (2) through (4) of Table 1.5 correspond to the three bonds chosen from Table 1.2 to construct the replicating portfolio. Row (iii) gives the face amount of each bond in the replicating portfolio, that is, the portfolio is long about 2.90 face amount of the 2.875s of 11/15/2021; about 2.94 of the 2.125s of 05/15/2022; and about 102.98 of the 1.625s of 11/15/2022. Rows (iv) through (vi) show the cash flows from those face amounts of each bond. For example, 102.97582 face amount of the 1.625s of 11/15/2022 generates coupon payments of $102.97582 \times 1.625\%/2 = 0.83668$ on November

[4]Market participants often use the term arbitrage more broadly to encompass trades that could conceivably lose money but promise large profits relative to risks.

[5]Shorting a bond means selling a bond one does not own. The mechanics of short selling bonds is discussed in Chapter 10. For now, assume that a trader shorting a bond receives the price of the bond at the time of sale and is obliged to pay all of its coupons and its principal amount on their respective due dates.

TABLE 1.5 An Arbitrage Trade: Selling the 7.625s of 11/15/2022 and Buying a Replicating Portfolio from Bonds in Table 1.2. Coupons Are in Percent.

	(1)	(2)	(3)	(4)	(5)
			Replicating Portfolio		
(i)	Coupon	2.875	2.125	1.625	7.625
(ii)	Maturity	11/15/2021	05/15/2022	11/15/2022	11/15/2022
(iii)	Face	2.90281	2.94454	102.97582	100
	Date		**Cash Flows**		
(iv)	11/15/2021	2.94454	0.03129	0.83668	3.8125
(v)	05/15/2022		2.97583	0.83668	3.8125
(vi)	11/15/2022			103.8125	103.8125
(vii)	Price	101.4297	102.0662	102.2862	111.3969
(viii)	Cost	2.94431	3.00538	105.33005	111.3969
(ix)	Total Cost		111.2797		111.3969
(x)	Net Proceeds		111.3969 – 111.2797 = 0.1172		

15, 2021, and on May 15, 2022, and coupon plus principal payments of $102.97582 \times (1 + 1.625\%/2) = 103.8125$ on November 15, 2022. Row (vii) gives the price of each of the bonds, simply copied from Table 1.2, and row (viii) gives the cost of purchasing the face amount of each bond given in row (iii). As an example of the latter, the cost of purchasing 2.94454 face amount of the 2.125s of 05/15/2022 at a price of 102.0662 (per 100 face amount) is $2.94454 \times 102.0662\% = 3.00538$.

Column (5) of Table 1.5 gives details of the rich bond to be sold, namely, the 7.625s of 11/15/2022. Its cash flows on the three payment dates are given in rows (iv) through (vi). And, most importantly, note that the sums of the cash flows of the three replicating bonds for each date equal the cash flows of the 7.625s of 11/15/2022. More specifically, $2.94454 + 0.03129 + 0.83668 = 3.8125$; $2.97583 + 0.83668 = 3.8125$; and trivially, $103.8125 = 103.8125$. Hence, the replicating portfolio, as defined by the three bonds and their respective face amounts in row (iii), do indeed replicate the cash flows of the 7.625s of 11/15/2022. Section A1.1 in the appendix to this chapter shows how to derive the face amounts of the bonds in this or any such replicating portfolio.

The discussion now returns to the arbitrage trade. According to row (ix) of Table 1.5, an arbitrageur can sell 100 face amount of the 7.625s of 11/15/2022 for 111.3969 and buy the replicating portfolio for 111.2797, which is just the sum of the costs of the three bond positions given in row (viii). The net proceeds of this trade, given in row (x), is 0.1172, or about

12 cents. But, by definition of the replicating portfolio, this trade will neither generate nor require cash on any of the three future payment dates. Hence, through this trade, the arbitrageur receives 12 cents today without incurring any future obligations. While these proceeds may seem small, the trade described in Table 1.5 can, at least in theory, be scaled up dramatically. For example, selling $500 million of the 7.625s of 11/15/2022 and buying the appropriately sized replicating portfolio would generate a riskless profit of $\$500{,}000{,}000 \times 0.1172\% = \$586{,}000$.

As discussed at the start of this section, if a riskless and profitable trade like the one just described were readily available, arbitrageurs would collectively rush to do the trade and, in so doing, force prices to relative levels that admit no further arbitrage opportunities. In the present example, arbitrageurs would drive the price of the 7.625s of 11/15/2022 lower and the price of the replicating portfolio higher until the two were equal.

The crucial link between arbitrage and the law of one price can now be explained. The total cost of the replicating portfolio, 111.2797, given in row (ix) of Table 1.5, exactly equals the present value of the 7.625s of 11/15/2022 as reported in Table 1.4. In other words, exactly the same value for the 7.625s of 11/15/2022 is computed through the law of one price (i.e., applying discount factors derived from the prices of bonds in the replicating portfolio) and through arbitrage pricing (i.e., finding the price of the replicating portfolio). This is not a coincidence. Section A1.2 in the appendix proves that these two pricing methodologies are mathematically identical. Hence, applying the law of one price, that is, pricing with discount factors, is identical to relying on the activity of arbitrageurs to eliminate relative mispricings. Put another way, discounting cash flows can be justifiably regarded as shorthand for the more complex and persuasive arbitrage pricing methodology.

Despite the theory, of course, as of mid-May 2021, the market price of the 7.625s of 11/15/2022 is quoted at about 12 cents more than the value of the replicating portfolio. This can be attributed to one or a combination of factors. First, there are transaction costs in doing arbitrage trades that could significantly lower or even wipe out any arbitrage profit. In particular, all of the prices in this chapter are ask prices, whereas, in reality, an arbitrageur might have to buy bonds at ask prices but sell other bonds at lower bid prices. Second, transaction costs in the financing markets (see Chapter 10), which are incurred when borrowing money to buy bonds and when borrowing bonds to short them, might similarly overwhelm any arbitrage profit. Third, it is only in theory that US Treasury bonds are commodities, that is, fungible collections of cash flows. In reality, bonds have idiosyncratic characteristics that manifest themselves in prices. The discussion of Table 1.4 mentions the richness of older, high-coupon bonds, and the next section will show idiosyncratic differences in the pricing of Treasury STRIPS.

1.5 APPLICATION: IDIOSYNCRATIC PRICING OF TREASURY STRIPS

In contrast to coupon bonds that make payments every six months, *zero coupon* bonds make no payments until maturity. Zero coupon bonds issued by the US Treasury are called STRIPS. For example, $1,000,000 face amount of STRIPS maturing on May 15, 2030, promises only one payment: $1,000,000 on that date. STRIPS are created when a particular coupon bond is delivered to the Treasury in exchange for claims on its future coupon and principal components. Table 1.6 illustrates the stripping of $1,000,000 face amount of the 0.625s of 05/15/2030. As of mid-May 2021, the bond has nine years remaining to maturity, which means it will make 18 coupon payments, from November 15, 2021, to May 15, 2030, and one principal payment, on May 15, 2030. STRIPS received in exchange for coupon (interest) payments are called TINTs, INTs, or C-STRIPS, while STRIPS received in exchange for principal payments are called TPs, Ps, or P-STRIPS. Table 1.6 shows that stripping $1,000,000 face amount of the 0.625s of 05/15/2030 generates $\$1{,}000{,}000 \times 0.625\%/2 = \$3{,}125$ face amount of C-STRIPS maturing on each coupon payment date and $1,000,000 face amount of P-STRIPS maturing on the bond's maturity date.

The Treasury not only creates STRIPS but retires them as well. For example, upon delivery of all of the STRIPS in Table 1.6, the Treasury would *reconstitute* the $1,000,000 face amount of the 0.625s of 05/15/2030. It is critical to note, however, that C-STRIPS are fungible while P-STRIPS are not. When reconstituting a bond, any C-STRIPS maturing on a particular date may be applied toward the coupon payment of that bond on that date. By contrast, only P-STRIPS that were stripped from a particular bond may

TABLE 1.6 STRIPS Created from $1 Million Face Amount of the 0.625s of 05/15/2030, as of Mid-May 2021. Entries Are in Dollars.

Date	C-STRIP Face Amount	P-STRIP Face Amount
11/15/2021	3,125	
05/15/2022	3,125	
11/15/2022	3,125	
...	...	
05/15/2029	3,125	
11/15/2029	3,125	
05/15/2030	3,125	1,000,000

be used to reconstitute the principal payment of that bond.[6] This feature of the STRIPS program implies that P-STRIPS, and not C-STRIPS, are likely to inherit the cheapness or richness of the bonds from which they are stripped.

STRIPS prices are essentially discount factors. If the price of C-STRIPS maturing on May 15, 2030, is 85.9453 per 100 face amount, then the implied discount factor to that date is 0.859453.

As of mid-May 2021, the STRIPS market provides another illustration of the idiosyncratic pricing of US Treasury bonds, that is, of the occasional failures of the law of one price to describe prices accurately. Table 1.7 isolates one such case by reporting the prices of three STRIPS that all mature on May 15, 2030. In theory, because each of these zero coupon bonds pay $1 on May 15, 2030, they should all sell for the same price. But they don't. The C-STRIPS maturing on that date are priced at 85.95, while two P-STRIPS, one from stripping the 0.625s of 05/15/2030 and the other from stripping the 6.25s of 05/15/2030, are priced at 86.85 and 87.04, respectively. Reiterating the discussion earlier, these price discrepancies do not necessarily imply an arbitrage opportunity of buying the C-STRIPS and shorting one of the P-STRIPS because of likely significant market frictions. An investor planning to buy and hold one of these zero coupon bonds to maturity, however, would certainly be most attracted to the C-STRIPS.

Figure 1.2 shows a broader pattern of the idiosyncratic pricing of STRIPS. The thick, round data points are C-STRIPS prices, and the plus signs are P-STRIPS prices, where prices are on the left axis. As an aside, because STRIPS prices are discount factors times 100, the C-STRIPS data points depict a collection of discount factors out to 30 years, down to a value of 48.60 at 30 years. But returning to the main point, many P-STRIPS prices are clearly not identical to C-STRIPS prices of identical maturity. To highlight this point, the difference between P-STRIPS and C-STRIPS

TABLE 1.7 Prices of STRIPS Maturing on May 15, 2030, as of May 14, 2021.

STRIPS	Price
C-STRIPS	85.9453
P-STRIPS of 0.625s of 5/15/2030	86.8549
P-STRIPS of 6.25s of 5/15/2030	87.0412

[6] Making P-STRIPS fungible with C-STRIPS would not affect either the total or the timing of cash payments by the Treasury but could change the amount outstanding of particular bonds.

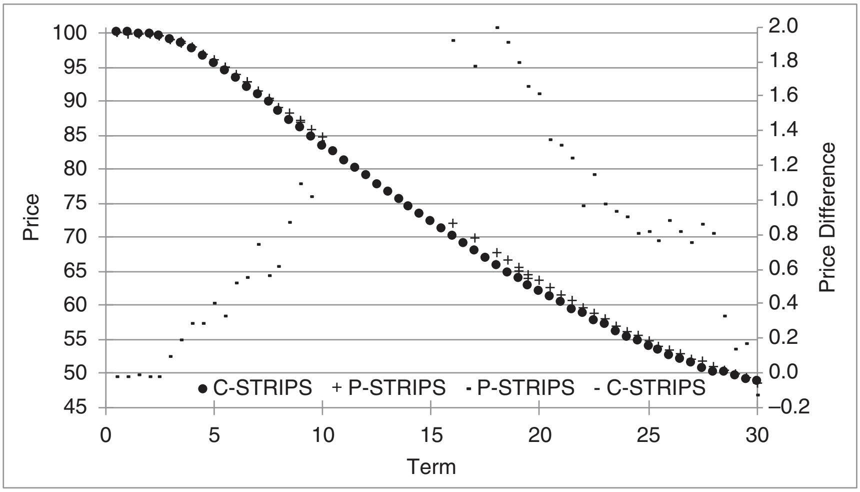

FIGURE 1.2 Prices of US Treasury C-STRIPS and P-STRIPS, as of May 14, 2021.

prices of the same maturity are graphed as short dashes against the right axis. A few of the differences are negative, but most are positive, and some significantly so, reaching a peak of 1.91 per 100 face amount at a term of 18.5 years.

1.6 ACCRUED INTEREST

This section describes the market practice of separating the full or invoice price of a bond, which is the price paid by a buyer to the seller, into two parts: a flat or quoted price, which appears on trading screens and is used when negotiating transactions; and accrued interest. Full and flat prices are also known as *dirty* and *clean* prices, respectively.

For concreteness, consider an investor who purchases \$10,000 face amount of the US Treasury 0.625s of 08/15/2030, for settlement on May 17, 2021. The bond last made a coupon payment of $\$10{,}000 \times 0.625\%/2 = \$31.25$ on February 15, 2021, and makes its next coupon payment of \$31.25 on August 15, 2021. See the timeline in Figure 1.3.

Assuming that the buyer holds the bond through August 15, 2021, the buyer collects the semiannual coupon on that date. But it can be argued that the buyer is not really entitled to the whole coupon because the buyer will have held the bond for only three months, that is, from May 17, 2021, to August 15, 2021. More precisely, using what is known as the *actual/actual day-count convention*, and referring again to Figure 1.3, the buyer should receive only 90 of the 181 days of the coupon payment, that is,

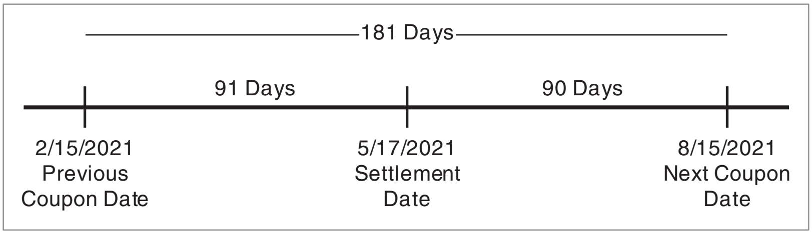

FIGURE 1.3 Timeline for Computing the Accrued Interest on the 0.625s of 08/15/2030 for Settlement on May 17, 2021.

$\$31.25 \times 90/181 = \$15.539$. The seller, on the other hand, who presumably held the bond from the previous coupon date to the settlement date, should collect the rest of the coupon, namely, the accrued interest from February 15, 2021, to May 17, 2021, which is $\$31.25 \times 91/181 = \$15.711$.

A conceivable institutional arrangement would be for the seller and the buyer to divide the coupon payment on the next payment date, but this would undesirably require additional arrangements to ensure compliance. Instead, market convention dictates that the buyer pay the seller the $15.711 of accrued interest on the settlement date, and that the buyer keep the entire coupon of $31.25 on the coupon payment date.

The flat or quoted price of the 0.625s of 08/15/2030 on May 14, 2021, for settlement on May 17, is 91.78125. The full or invoice price, which is defined as the flat or quoted price plus accrued interest, is $91.78125 + 0.15711 = 91.93836$. For the particular trade just described, of $10,000 face amount, the invoice amount is $\$10{,}000 \times (91.78125\% + 0.15711\%) = \$9{,}178.125 + \$15.711 = \$9{,}193.836$, or, equivalently, $\$10{,}000 \times 91.93836\% = \$9{,}193.836$.

At this point, it becomes clear why earlier in the chapter it is noted that prices are full prices. With bonds paying coupons on May 15 and November 15, and with trades settling on Monday, May 17, 2021, buyers have to pay sellers two days of accrued interest.

The full price of a bond, the amount a buyer actually pays to purchase a bond, should equal the present value of a bond's cash flows. Mathematically, denote the flat price of the bond by p, accrued interest by AI, the present value of the cash flows by PV, and the full price by P. Then,

$$P = p + AI = PV \tag{1.5}$$

Equation (1.5) reveals that the particular market convention used to calculate accrued interest does not really matter. For example, the accrued interest convention just described is conceptually too generous to the seller because, instead of being made to wait until the next coupon date, the seller receives a share of the coupon at settlement. But the accrued interest convention changes neither the present value of all cash flows, PV, in Equation (1.5), nor the full price, P, which is the amount actually exchanged at settlement. The convention only changes the quoted market price, p, relative to the calculated amount of accrued interest. If accrued interest is in any sense too high, the market reduces p accordingly.

Having made this argument, why bother with the accrued interest convention in the first place; that is, why not just quote bonds using the full price? The answer is told in Figure 1.4, which draws the full and flat prices of the 0.625s of 08/15/2030 from February 15, 2021, to September 15, 2021, under the simplifying assumption that interest rates are constant at the approximate market rate of 1.60% over the entire period.[7]

Figure 1.4 shows that the full price changes dramatically over time – including a sudden drop on August 15, 2021 – even though the market, in terms of interest rates, is unchanged. The flat price, by contrast, changes only gradually over time. Therefore, when trading bonds day to day, it is more intuitive to follow flat prices and negotiate transactions in those terms.

The behavior of the prices in Figure 1.4 can be understood readily. First, as of February 2021, the price of the bond is below 100, because its coupon rate is below the market rate for its maturity. But as the bond matures, its flat price approaches 100. (The dynamics of prices at below and above market

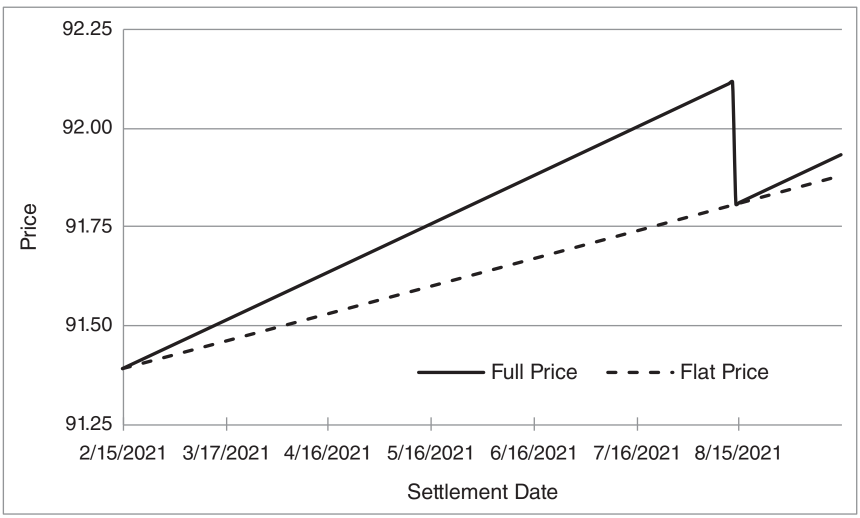

FIGURE 1.4 Full and Flat Prices for the 0.625s of 08/15/2030, Assuming Constant Interest Rates.

[7]More specifically, it is assumed that the continuously compounded interest rate for all terms is constant at 1.60%.

rates are discussed further in Chapter 3.) Second, within a coupon period, the full price of a bond increases over time as its cash flows get closer to being paid, that is, as their present values increase. But from just before the coupon payment date to just after, the full price falls by the coupon payment: the coupon is included in the present value just before the payment, but not included just after. The flat price, by contrast, which equals the full price minus accrued interest, rises more gradually than the full price and does not fall precipitously just after the coupon payment. Because accrued interest just before the coupon payment nearly equals that coupon payment, subtracting accrued interest from the full price leaves the flat price essentially without that coupon – both just before and just after its payment date.

1.7 DAY-COUNT CONVENTIONS

As mentioned in the previous section, accrued interest is calculated using the actual/actual convention, that is, by dividing the actual number of days from the last coupon payment by the actual number of days between coupon payments. Hence the term "actual/actual" for this day-count convention.

The actual/actual convention is commonly used in government bond markets, but, in other markets, different conventions are used. Two of the most common are *actual/360* and *30/360*. The actual/360 convention divides the actual number of days between two dates by 360. The 30/360 convention calculates the difference between two dates under the assumption that there are 30 days in each month, and then divides that difference by 360. To illustrate, there are 75 actual days between June 1 and August 15: there are 29 days from June 1 to the end of June; 31 days in July; and 15 days to August 15. The 30/360 convention, however, assumes that there are 30 days in every month, including July, giving a total of 29 + 30 + 15 or 74 days. Money markets typically use the actual/360 convention; swap markets use either actual/360 or 30/360; and corporate bond markets typically use the 30/360 convention.

CHAPTER 2

Swap, Spot, and Forward Rates

Chapter 1 showed that discount factors fully describe the time value of money as embedded in market prices. Investors and traders, however, often find it more intuitive to quote the time value of money in terms of interest rates and, in particular, in terms of either *swap rates* or *par rates*, *spot rates*, and *forward rates*.

This chapter begins by explaining that interest rates are always quoted as annual rates, that interest is conceptualized as being paid over a number of periods of fixed length (e.g., 90 days, three semiannual periods), and that interest rate quotations indicate the payment of either *simple* or *compound* interest. The chapter then introduces interest rate swaps (IRS) as context for the material. Swaps and bonds together comprise a significant portion of fixed income markets, and swaps, because they are relatively liquid, have become benchmarks against which to evaluate other fixed income instruments.

At the time of this writing, interest rate swap markets are in transition away from LIBOR (London Interbank Offered Rate), which has dominated floating-rate indexes for decades. Chapter 12 discusses this transition, but this chapter briefly introduces the leading candidates for replacing LIBOR (e.g., *Secured Overnight Financing Rate (SOFR)* in the United States) and their associated swaps. Chapter 13 discusses why and how market participants use interest rate swaps.

2.1 INTEREST RATE QUOTATIONS

An investment with fixed cash flows is completely described by its price and cash flows. It seems enough to know, for example, that a bond paying 102 in six months costs 101.4925 today, or that a $100,000,000 1.5-year loan made six months from today will return $103,030,100 in two years. Nevertheless, investors and traders often prefer to quote and think about valuations in terms of interest rates. As shown herein, the bond and loan just described earn semiannually compounded rates of 1% and 2% per year, respectively. Rates are more intuitive than prices and cash flows because rates automatically normalize for the amount invested and for the investment horizon. The bond costs 101.4925 and matures in six months, while the loan invests $100,000,000 in six months for 1.5 years, but their respective interest rates can be sensibly and intuitively compared.

Interest rates are always quoted as annual rates over a *term* or *tenor*, which is described as a number of periods of fixed length. A money market instrument, for example, might offer 1% per year for 90 days on an investment of $100,000. If this 1% is quoted as a simple rate of interest, under the actual/360 convention described in Section 1.7, then the interest payment at the end of the 90 days is,

$$\$100{,}000 \times 1\% \times \frac{90}{360} = \$250.00 \tag{2.1}$$

If the 1% interest rate on the money market instrument is quoted not as a simple rate, but as a *daily compounded* rate, then interest is computed differently: the investment earns simple interest daily, but interest is earned on accumulated balances, which include interest already earned. After the first day, then, the balance includes one day of simple interest,

$$\$100{,}000\left(1+\frac{1\%}{360}\right) = \$100{,}002.78 \tag{2.2}$$

Over the second day, the entire balance at the end of the first day, given in Equation (2.2), earns another day of simple interest. This implies that interest is compounded, that is, the $2.78 of interest earned over the first day itself earns interest on the second day. As a result, the total balance at the end of the second day is,

$$\$100{,}002.78\left(1+\frac{1\%}{360}\right) = \$100{,}000\left(1+\frac{1\%}{360}\right)^2 = \$100{,}005.56 \tag{2.3}$$

where the first equality comes from substituting the left-hand side of (2.2) for the $100,002.78 in (2.3).

Continuing in this fashion, the total balance at the end of the 90 days, which is the amount paid by the money market instrument at maturity, is,

$$\$100{,}000\left(1+\frac{1\%}{360}\right)^{90} = \$100{,}250.31 \tag{2.4}$$

Comparing the interest component of Equation (2.4), that is, $250.31, with that of (2.1), that is, $250.00, shows the effect of compound interest in this example. A daily compounded rate of 1% on $100,000 earns 31 cents more than a simple interest rate of 1%. When rates are higher, and when the term of the investment is longer, the difference can be much larger.

Return now to the bond and loan examples given at the start of this section. In these cases, interest is likely to be semiannually compounded, where, by convention, a semiannual period is exactly one half of a year. Therefore, the six-month bond, which costs 101.4925 today and pays 102 after six months, earns 1% in the sense that,

$$\$101.4925\left(1+\frac{1\%}{2}\right)=102 \tag{2.5}$$

The loan invests \$100 million six months from today and, three semiannual periods (i.e., 1.5 years) later, returns \$103,030,000. This investment earns a semiannually compounded rate of 2% in the sense that,

$$\$100{,}000{,}000\left(1+\frac{2\%}{2}\right)^3=\$103{,}030{,}100 \tag{2.6}$$

This section illustrates concepts with daily and semiannual periods. Daily periods are common in money markets and in the swap market discussed in the next section. Semiannual periods are common in many government and corporate bond markets, like the US government bond market, discussed in Chapter 1. Different periods are used in other contexts, however. Mortgage markets use monthly periods, for example, because mortgage payments are typically paid monthly.

More generally, then, let n be the number of periods per year; let N be the number of periods; and let $\hat{r}$ be the interest rate to be compounded n times per year. Then, investing F at the rate $\hat{r}$ grows, after N periods, to a total balance of,

$$F\left(1+\frac{\hat{r}}{n}\right)^N \tag{2.7}$$

Markets should ensure that the final proceeds from identical investments of the same term are the same, regardless of the compounding convention. If, for example, the market offers 102 on a 1-year investment of 100, that investment might be quoted as earning 2% annually; 1.9901% semiannually compounded; or 1.9819% compounded monthly, because,

$$100(1+2\%)=100\left(1+\frac{1.9901\%}{2}\right)^2=100\left(1+\frac{1.9819\%}{12}\right)^{12}=102 \tag{2.8}$$

Note that the greater the frequency of compounding, the lower the quoted interest rate. These rates are set such that more frequent compounding, that is, paying more interest on interest, exactly offsets the lower rate earned on the initial investment.

This section concludes by noting that some contexts use *continuous compounding*, in which interest is conceptualized as being paid every instant. This quoting convention can appear in money markets; is frequently used at financial firms, because rates and discount factors have to be calculated at irregular intervals; and is used almost exclusively by researchers for mathematical convenience. Continuous compounding is presented in Appendix A2.1.

2.2 INTEREST RATE SWAPS

In an interest rate swap, two parties agree to exchange a series of interest payments. Consider first the solid arrows in Figure 2.1, which are the only contractual cash flows of the depicted swap. Counterparty A agrees to pay Counterparty B the fixed *swap rate* of 0.1120% annually for two years on a notional amount of $100 million. In return, Counterparty B agrees to pay to Counterparty A daily compounded SOFR annually for two years on this same notional amount. No cash is exchanged on the trade or settlement date. Counterparty A is said to *pay fixed* and *receive floating*, while Counterparty B is said to *receive fixed* and *pay floating*.

As is explained in Chapter 12, SOFR on any given day is the volume-weighted median rate at which market participants borrow and lend overnight funds secured by US Treasury collateral. Conceptually, daily SOFR represents the rate on extremely safe, overnight loans made that day.

The $100 million in the swap of Figure 2.1 is called the *notional amount* of the swap, rather than the face, par, or principal amount of the swap, because it is used only to compute the fixed- and floating-rate payments. The notional amount is never paid or received by either counterparty. The dashed arrows in Figure 2.1, which portray a final exchange of notional amount, will be convenient later for pricing the swap, but this exchange is fictional: it is not part of the swap contract and never actually takes place.

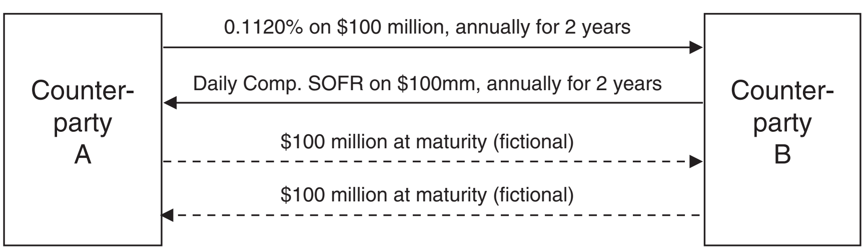

FIGURE 2.1 A SOFR Swap.

Payments on both sides of a SOFR swap follow the actual/360 day-count convention. The annual interest payment on the fixed leg of the swap, therefore, over any 365-day year, is,

$$\$100{,}000{,}000 \times 0.1120\% \times \frac{365}{360} = \$113{,}556 \tag{2.9}$$

Over 366-day years, of course, the fraction in Equation (2.9) would be 366/360 instead.

The annual payment on the floating side can be computed only after every daily SOFR rate that year has been realized. To illustrate, assume a very simple scenario in which, over a 365-day year, SOFR was 0.10% for five days; 0.50% for 170 days; and 0.01% for 190 days. In that scenario, an investment of \$100 million compounded daily at SOFR rates would grow to,

$$\$100{,}000{,}000\left(1+\frac{0.10\%}{360}\right)^{5}\left(1+\frac{0.50\%}{360}\right)^{170}\left(1+\frac{0.01\%}{360}\right)^{190}$$
$$= \$100{,}243{,}071 \tag{2.10}$$

Therefore, the payment on the floating leg at the end of the year, representing the interest on a daily compounded SOFR investment over that year, is,

$$\$100{,}243{,}071 - \$100{,}000{,}000 = \$243{,}071 \tag{2.11}$$

In this particular scenario, Counterparty A receives \$243,071 on the floating leg, but pays only \$113,556 on the fixed leg. Had realizations of SOFR over the year been lower, the reverse could have easily been true, with Counterparty A receiving less on the floating leg than it paid on the fixed leg.

SOFR swaps that mature in an exact number of years, like the example in Figure 2.1, make annual payments. SOFR swaps that mature in less than one year make one payment at maturity. And SOFR swaps that mature in more than one year, but not in an exact number of years, typically make one *stub* payment followed by annual payments to maturity. A 1.5-year swap, for example, would likely make a stub payment after six months and another payment one year later.

Figure 2.2 shows the *term structure* of swap rates in different currencies as of mid-May 2021. The SOFR curve, for example, gives the fixed rates that can be exchanged for SOFR for various terms. As highlighted in Figure 2.1, in a two-year US dollar (USD) swap, 0.1120% fixed can be exchanged for floating SOFR. Figure 2.2 shows additional USD swap rates, for example, 1.352% for a 10-year swap and 1.758% for 30 years.

The other rates and currencies shown are SONIA (Sterling Overnight Interbank Average) in British Pounds (GBP); TONAR (Tokyo Overnight

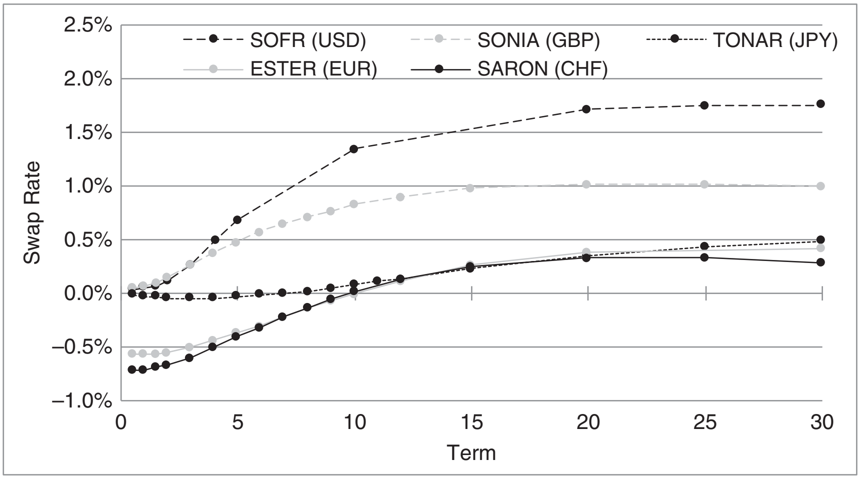

FIGURE 2.2 Term Structures in Different Currencies, as of May 14, 2021.

Average Rate) in Japanese Yen (JPY); ESTER, also written as €STR (Euro Short-Term Rate) in Euro (EUR); and SARON (Swiss Average Rate Overnight) in Swiss Franc (CHF). These rates, which are discussed further in Chapter 12, are often called "risk-free" rates to distinguish them from LIBOR rates, but only SOFR and SARON are rates on loans collateralized by government bonds. SONIA, TONAR, and ESTER, by contrast, are all rates on interbank loans. In any case, Figure 2.2 illustrates that the term structures of these rates vary across currencies and, of course, though not shown, over time as well.

2.3 PRICING INTEREST RATE SWAPS

As mentioned before, it is convenient for pricing purposes to assume the fictional exchange of notional amounts depicted in Figure 2.1. First, an exchange of notional amount does not change the value of the swap to either counterparty: paying \$100 million and receiving \$100 million at exactly the same moment has no value. Second, including the receipt of the fictional notional amount, Counterparty B's position very much resembles buying a fixed-rate bond: B receives an annual coupon and then, at maturity, receives a final "principal" payment. Third, including the receipt of the fictional notional amount, Counterparty A's position very much resembles buying a floating-rate bond, receiving interest that depends on interest rates as they evolve, and then receiving a final "principal" payment.

The value of a floating-rate bond that always pays the fair market rate is worth par, or face amount, today. In terms of Figure 2.1, the value to

Counterparty A of receiving floating, including the fictional notional amount, is \$100,000,000. Intuitively, by the definition of SOFR as a money market rate, Counterparty A can lend \$1 to a dealer on any day and the next day collect that \$1 plus SOFR interest. Therefore, Counterparty A and the dealer would similarly be willing to exchange \$100,000,000 today for a floating-rate bond that pays compounded SOFR over some term and \$100,000,000 at the end of that term.

With the floating leg of the swap – including the fictional notional amount – worth par, Counterparty B's position can be recast as buying a two-year, 0.1120% bond for par, that is, for \$100,000,000, where par happens to take the form of the floating leg of the swap. In terms of the bond pricing methodology of Chapter 1, the present value of the payments on the fixed leg of the swap, including the fictional notional amount, must equal par. This insight is so useful and commonplace that the phrase, "fixed leg of a swap," is almost always meant to include the fictional notional amount, and the text will adopt this convention from here on.

Counterparty B's position can also be recast as buying a two-year, 0.1120% for par, and financing that position by borrowing at the floating rate. To elaborate, the cash flows from receiving fixed in swap very much resemble those from a leveraged long position in a bond. At initiation, Counterparty B neither receives nor pays cash: the swap has no initial cash flow, and the bond is purchased with borrowed money. Over the life of the trade, Counterparty B receives a fixed rate and pays a floating rate: fixed minus floating on the swap, and the coupon rate minus the short-term financing rate on the bond. And, at maturity, Counterparty B neither receives nor pays a final amount: the swap does not exchange notional amount, and the principal of the bond is used to repay the original loan. In a similar fashion, Counterparty A's position can be recast as borrowing a bond to short it. This perspective on swaps is very useful for understanding how market participants use swaps to adjust their interest rate exposures, which is discussed in Chapter 13.

Given the discussion to this point, it is clear that swap rates can also be called *par rates*. A par rate of a certain term is defined as the rate paid on an investment that costs par and promises to repay par at maturity. If, for example, a 10-year US Treasury bond has a coupon of 1.625% and costs 100, then 1.625% is the 10-year par rate in the Treasury market. While, in actuality, there is neither "investment" in a swap nor a repayment of par, thinking of a swap with an exchange of fictional notional amounts allows for the interpretation of the swap rate as the rate "earned" on the fixed leg, that is, on a par investment.

The text now turns to applying the pricing insights of this section to extract discount factors from SOFR swap rates. To illustrate, the first and second columns of Table 2.1 list SOFR swap rates of terms 0.5 through 2.0 years, as of May 14, 2021, and Table 2.2 lists the payment dates of these

TABLE 2.1 Swap Rates, Spot Rates, and Forward Rates Implied by USD SOFR Swaps, as of May 14, 2021. Rates Are in Percent.

Term	Swap Rate	Spot Rate	Forward Rate	Discount Factor
0.5	0.0340	0.0348	0.0348	0.999826
1.0	0.0460	0.0466	0.0585	0.999534
1.5	0.0670	0.0681	0.1111	0.998979
2.0	0.1120	0.1136	0.2500	0.997732

TABLE 2.2 Days from Settlement or Previous Payment Date for SOFR Swaps Settling on May 18, 2021.

Term	11/18/2021	05/18/2022	11/18/2022	05/18/2023
0.5	184			
1.0		365		
1.5	184		365	
2.0		365		365

swaps and the number of days between payments. (Note that prices are as of Friday, May 14, and settlement occurs two business days later, on Tuesday, May 18.) Using these tables, the following equations link discount factors to quoted SOFR swap rates,

$$100\left(1+0.0340\%\frac{184}{360}\right)d(0.5)=100 \tag{2.12}$$

$$100\left(1+0.0460\%\frac{365}{360}\right)d(1)=100 \tag{2.13}$$

$$0.0670\frac{184}{360}d(0.5)+100\left(1+0.0670\%\frac{365}{360}\right)d(1.5)=100 \tag{2.14}$$

$$0.1120\frac{365}{360}d(1)+100\left(1+0.1120\%\frac{365}{360}\right)d(2)=100 \tag{2.15}$$

In words, Equations (2.12) through (2.15) say that the present value of the fixed leg of each SOFR swap equals par. Equation (2.12) says that, for 100 notional amount of the 0.5-year swap, which matures in 184 days, the fixed leg pays 100 plus interest on 100 at the 0.5-year swap rate of 0.0340%. Furthermore, that payment times the six-month discount factor equals par, or 100. Equation (2.13) says the same for the one-year swap, with its single payment, at the one-year swap rate of 0.0460%, payable in 365 days.

The fixed leg of the 1.5-year swap makes two payments. Following the rule explained in the previous section, the first payment is made in six months so that the subsequent payment can be made exactly one year later. According to Tables 2.1 and 2.2, the payment in six months, in this case 184 days, on 100 face amount of the 1.5-year swap, is 0.0670 times 184/360. The payment one year later, which is 1.5 years from settlement, includes interest over that 365-day year plus the notional amount, for a total of 100 plus 0.0670 times 365/360. Equation (2.14) takes the present value of these two payments, multiplying the first by $d(0.5)$ and the second by $d(1.5)$, and sets the sum equal to par.

Lastly, the fixed leg of the two-year swap makes two payments, the first in 365 days and the next 365 days later. Each interest payment, therefore, on 100 face amount, is 0.1120 times 365/360, and the notional amount is assumed paid at the end of the second year. Hence, Equation (2.15) sets the present value of the fixed leg of this swap equal to par.

Solving Equations (2.12) through (2.15) for the unknown discount factors, analogously to the solution for discount factors in the US Treasury market in Section 1.2, gives the discount factors in Table 2.1. Chapter 1 introduced discount factors as an expression of the time value of money, and this chapter, so far, has added swap or par rates. The next two sections continue with other popular expressions of the time value of money, namely spot and forward rates.

2.4 SPOT RATES

The word *spot* in finance typically refers to transactions for immediate or imminent settlement, as opposed to *forward* transactions, which settle further in the future. Consistent with this usage, a spot rate is the rate on a spot loan, an agreement in which a lender gives money to the borrower at or around the time of the agreement and, furthermore, expects repayment at some single, specified time in the future. For example, along the lines of Equation (2.7), a two-year investment of 100, at a semiannually compounded spot rate of 0.1136%, grows over those two years or four semiannual periods to a final payment of,

$$100\left(1+\frac{0.1136\%}{2}\right)^{2\times 2} = 100.2274 \tag{2.16}$$

More generally, denote the semiannually compounded t-year spot rate by $\hat{r}(t)$. With semiannual compounding and an investment period of t years, or $2t$ semiannual periods, investing one unit of currency from now to year t generates final proceeds of,

$$\left(1+\frac{\hat{r}(t)}{2}\right)^{2t} \tag{2.17}$$

To link spot rates and discount factors, note that if one unit of currency grows to the quantity in (2.17) in t years, then the present value of that quantity, by definition, is one. Using discount factors to compute that present value,

$$\left(1+\frac{\hat{r}(t)}{2}\right)^{2t} d(t) = 1 \tag{2.18}$$

or, solving for $d(t)$,

$$d(t) = \frac{1}{\left(1+\frac{\hat{r}(t)}{2}\right)^{2t}} \tag{2.19}$$

Either of these two equations can be used to solve for a spot rate of term t given the discount factor of that term. To illustrate, consider the discount factors implied by SOFR swap rates in Table 2.1. The two-year discount factor is 0.997732, which, by either (2.18) or (2.19), implies that the two-year spot rate is 0.1136%. Along the same lines, Table 2.1 computes spot rates of terms 0.5 to 2.0 years from the respective discount factors.

2.5 FORWARD RATES

A forward rate is the rate on a forward loan, which is an agreement today to lend money at some time in the future and to be repaid some time after that. The agreement today specifies the forward rate, which means that the rate on the loan is set today even though the loan itself will not be made until a later date. There are many possible forward rates: the rate on a 1.5-year loan given six months from today; the rate on a six-month loan given in five years; etc. This section, however, focuses on forward rates over sequential, six-month periods. Let $f(t)$ denote the forward rate on a loan from year $t-0.5$ to year t. Then, investing one unit of currency from year $t-0.5$ for six months generates proceeds, at year t, of $1+f(t)/2$.

To link forward rates to spot rates, note that a spot loan from now to year $t-0.5$, combined with a forward loan from year $t-0.5$ to year t, covers the same investment period as a spot loan from now to year t. Consistent quoting of rates, therefore, ensures that the proceeds from these two alternatives are the same. Mathematically, noting that $t-0.5$ years is $2(t-0.5)$ or $2t-1$ semiannual periods,

$$\left(1+\frac{\hat{r}(t)}{2}\right)^{2t} = \left(1+\frac{\hat{r}(t-0.5)}{2}\right)^{2t-1}\left(1+\frac{f(t)}{2}\right) \tag{2.20}$$

Extending this logic to say that a spot loan to year t is equivalent to a series of six-month forward loans, spot rates and forward rates can also be linked as follows,

$$\left(1+\frac{\hat{r}(t)}{2}\right)^{2t}=\left(1+\frac{f(0.5)}{2}\right)\left(1+\frac{f(1.0)}{2}\right)\cdots\left(1+\frac{f(t)}{2}\right) \tag{2.21}$$

Note that $f(0.5)$, the rate on a "forward" loan from zero years to 0.5 years, is the same as the 0.5-year spot rate, $\hat{r}(0.5)$.

Forward rates can also be expressed in terms of discount factors. Use Equation (2.19) to substitute discount factors for spot rates in Equation (2.20) to see that,

$$1+\frac{f(t)}{2}=\frac{d(t-0.5)}{d(t)} \tag{2.22}$$

Applying the analytics of this section to the SOFR swaps in Table 2.1 allows for the computation of implied forward rates in that market. To illustrate with just one example, substitute the 1.5- and 2-year discount factors in the table into Equation (2.22) to derive the 2-year forward rate,

$$1+\frac{f(2)}{2}=\frac{0.998979}{0.997732} \tag{2.23}$$

$$f(2)=0.2500\% \tag{2.24}$$

2.6 RELATIONSHIPS BETWEEN SWAP, SPOT, AND FORWARD RATES

Table 2.1 gives the term structure of interest rates from the SOFR swap market in terms of swap, spot, and forward rates. This section describes several relationships between these curves that highlight their individual meanings. The discussion here is intuitive, while Appendix A2.2 takes a more mathematical approach.

The first relationship to be highlighted is that the t-year spot rate approximately equals the average of all the forward rates from today through year t. Taking one example from the table, the two-year spot rate is approximately equal to the average of the four forward rates from term 0.5 to 2 years,

$$\frac{0.0348\%+0.0585\%+0.1111\%+0.2500\%}{4}\approx 0.1136\% \tag{2.25}$$

This result makes sense because a two-year loan is equivalent to a combination of a six-month loan; a six-month loan six months forward; a six-month loan one year forward; and a six-month loan 1.5 years forward.

The second relationship, which is evident from Table 2.1, is that spot rates increase with term when forward rates are greater than spot rates. To take one specific example, the two-year spot, 0.1136%, is greater than the 1.5-year spot rate, 0.0681%, because the two-year forward rate, 0.2500%, is greater than the 1.5-year spot rate. Intuitively, it has just been established that spot rates are essentially an average of forward rates. But adding a number to an average increases that average if and only if the added number is larger than the previous average. Therefore, $\hat{r}(t) > \hat{r}(t-0.5)$ if and only if $f(t) > \hat{r}(t-0.5)$.

This relationship between spot and forward rates can also be seen from Figures 2.3 and 2.4. Figure 2.3 displays forward, spot, and swap curves in the SOFR market, and Figure 2.4 displays curves from the SARON market, which, as mentioned earlier, is a rate in Swiss Franc. The curves in both figures are based on market rates, as of May 14, 2021. SOFR rates are uniformly above SARON rates, which are actually negative over a significant range of terms.

The forward SOFR curve in Figure 2.3 increases to about the 12-year term and then declines, but it is always above the spot rate curve. Consistent with the previous paragraphs, the spot rate curve has to increase over the entire range of terms shown because forward rates are above spot rates. But the spot rate curve increases very mildly when forward rates are declining. True, spot rates have to continue to increase, but with forward rates exceeding spot rates by less and less, spot rates increase relatively slowly.

The forward SARON curve illustrates this relationship more dramatically. Until the 22-year term, forward rates exceed spot rates and, as

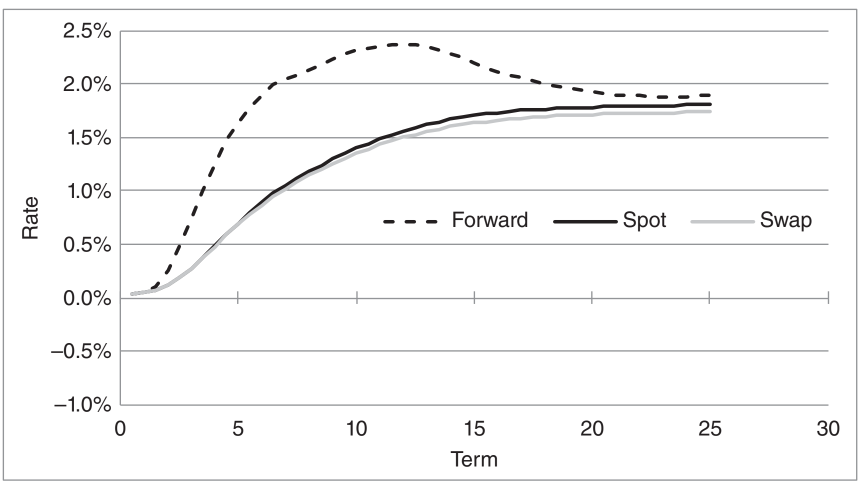

FIGURE 2.3 SOFR Rate Curves, as of May 14, 2021.

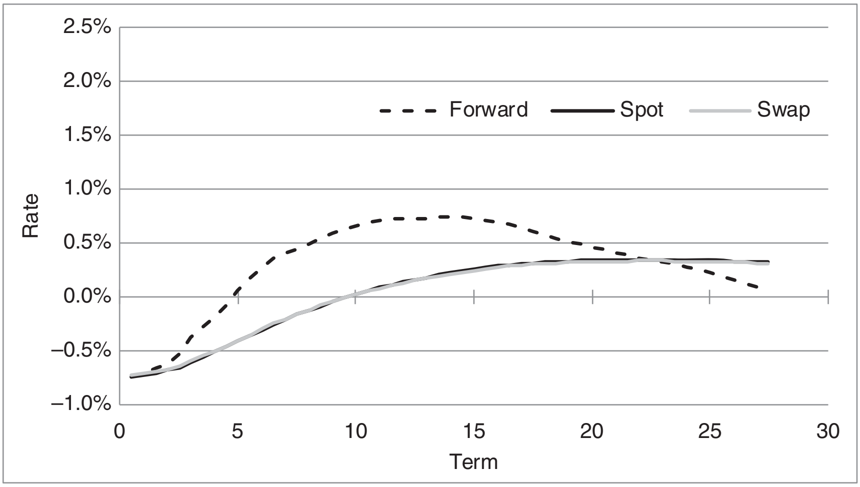

FIGURE 2.4 SARON Rate Curves, as of May 14, 2021.

expected, spot rates are increasing. From then on, however, forward rates are below spot rates and, although it is somewhat hard to see from the figure, spot rates begin to decline with term.

The third and final relationship between rates to be highlighted in this section is the following: when spot rates are increasing with term, swap or par rates are slightly below spot rates. This effect can be seen in both Figures 2.3 and 2.4, though the effect is more noticeable for the SOFR than for the SARON curves. To understand the intuition here, recall that the t-year spot rate is the return on an investment from today to year t, while the swap or par rate is the rate on an investment that pays every period from today to year t. The fair market swap rate, therefore, must reflect all of the spot rates from today to year t, though the t-year spot rate must be weighted most heavily as it is used to discount the by-far-and-away largest cash flow, that is, the fictional notional amount. When the term structure of spot rates is upward-sloping, then, all of the shorter-term spot rates are lower than the t-year spot rate, and the t-year par rate – which reflects all those spot rates – must be less than the t-year spot rate. But with the t-year spot rate weighted particularly heavily, the par rate is not very much below that spot rate.

CHAPTER 3

Returns, Yields, Spreads, and P&L Attribution

Chapter 2 showed how swap or par rates, spot rates, and forward rates are used to describe the interest rates that can be earned in various markets, like the Treasury bond market or the interest rate swap market. While these rates are derived from the prices of individual financial instruments, they are intended to describe the time value of money in a market as a whole. This chapter, by contrast, focuses on rates that are specific to individual bonds and then uses both market-wide and security-specific rates for *P&L (profit and loss) attribution*.

The first section of the chapter defines the realized return on a bond over a given horizon. *Ex-post* bond returns have to account for interim coupon payments and the reinvestment of those payments and are often computed on both a gross basis and net of financing, but are otherwise calculated like the returns on any other asset.

The next sections present *yield to maturity*. Bonds are often quoted and traded in terms of yield rather than price; yields are widely thought of as indicative of *ex-post* returns; and differences in bond yields are commonly regarded as indicative of differences in value. The discussion reveals, however, that yields can be misleading with respect to both *ex-post* returns and relative values.

The chapter continues with bond *spreads*. Because there are so many related but distinct fixed income products, spreads are used to quote, trade, and value one instrument relative to another. This part of the chapter describes various ways to compute spreads, explains their uses and limitations, and presents a detailed example of using spreads to assess the relative value of high-coupon US Treasury bonds.

The last sections of the chapter are devoted to attribution. In order to understand the performance of their trades and investments, traders and asset managers often divide their *ex-post* P&L or return into various components, for example, the passage of time; changing interest rates; and changing spreads. The text describes how to define the "passage of time," holding rates or spreads constant, and then illustrates P&L attribution with a detailed example of breaking into components the return on a particular US Treasury bond.

3.1 REALIZED RETURNS

The gross horizon return of a bond depends on the price at which the bond was bought; the coupon payments it earns over the horizon; the interest on the reinvestment of those coupon payments over the horizon; and the price of the bond at the end of the horizon. The net horizon return of the bond adjusts its gross return for the cost of financing its purchase.

Begin with a simple example. An investor buys \$1 million face amount of the US 7.625s of 11/15/2022 at 114.8765 in mid-November 2020. Six months later, in mid-May 2021, the price of the bond is 111.3969.[1] The gross return of the bond over the horizon is, therefore,

$$\frac{\$1{,}113{,}969 + \$38{,}125 - \$1{,}148{,}765}{1{,}148{,}765} = 0.2898\% \tag{3.1}$$

In words, the six-month holding period return in this example equals the value of the bonds at the end of the six months, or \$1,113,969; plus the coupon payment at that time of half 7.625% on the \$1 million face amount, or \$38,125; minus the initial cost of the bonds, or \$1,148,765, all divided by that initial cost. Note that the coupon income of the bond is sufficient to overcome the fall in price and result in a positive horizon return.

Realized return over a horizon that extends past a coupon payment depends on the rate at which coupon payments have been reinvested. If a bond makes a coupon payment in one month, but the horizon is two months, the horizon return depends on the one-month reinvestment rate, from the coupon payment date to the end of the horizon. For the purposes of illustration, consider the simple case of an investment in the 7.625s of 11/15/2022 over a horizon of one year. If the price of the bond at the end of the year is \$1,080,000, and if the coupon payment of \$38,125 after six months is reinvested for the subsequent six months at a money market rate of, say,

[1] For consistency with discussions later in the chapter, note that the two dates are actually November 13, 2020, and May 14, 2021. These two dates are Fridays, and the prices given here are full prices for settlement on November 16, 2020, and May 17, 2021, respectively, which means that the investor does not receive the coupon paid on November 15, 2020, but does receive the coupon paid on May 15, 2021.

0.05%, then the realized return of the investment over the year is,

$$\frac{\$1{,}080{,}000 + \$38{,}125\left(1 + \frac{0.05\%}{2}\right) + \$38{,}125 - \$1{,}148{,}765}{1{,}148{,}765}$$
$$= 0.6524\% \qquad (3.2)$$

The discussion now turns to the return net of the cost of financing the purchase of a bond. This might be an explicit cost; that is, a trader or investor might borrow the purchase price of the bond for six months at an interest rate of 0.05%. In that case, the interest cost of the borrowing over the investment horizon, $1,148,765 × 0.05%/2 = $287, would be subtracted from the numerator of Equation (3.1) to give a return net of financing of,

$$\frac{\$1{,}113{,}969 + \$38{,}125 - \$1{,}148{,}765 - \$287}{1{,}148{,}765} = 0.2648\% \qquad (3.3)$$

Not surprisingly, the net return of 0.2648% is the gross return, 0.2898%, minus the half-year borrowing cost of 0.05%/2, or 0.025%.

There are actually several subtleties in describing the net return on a bond investment in this way. First, market participants do fund bond purchases using the purchased bonds as collateral in the *repurchase* or *repo* market, which is the subject of Chapter 10. Second, only a portion of the purchase price can typically be funded in this market. In the present case, for example, an investor might be able to borrow only $1,125,790, or only 98%, of the total purchase price. In this case, the interest cost would only be half of 0.05% on that $1,125,790, or $281, and the net return would be 0.2653%, which is very slightly higher than the result from Equation (3.3). Third, regardless of how much is actually borrowed, the net return is calculated with the total purchase price as the denominator of (3.3). This formulation of net return, therefore, can be thought of as a return on balance sheet, that is, on the value that can be considered and reported as an asset. Fourth, return on capital would be computed differently. If a hedge fund borrows 98% of the value of the bond in the repo market, while putting up 2%, or $22,975 of its own capital, then this investment's return on capital – abstracting from any additional allocations of risk capital to the trade – would be,

$$\frac{\$1{,}113{,}969 + \$38{,}125 - \$1{,}148{,}765 - \$281}{\$22{,}975} = 13.27\% \qquad (3.4)$$

The 13.27% return on capital is 50 times the net balance sheet return of 0.2653%. This high return on capital is due to the fact that the trade, as described, has a leverage of 50, that is, an asset value of $1,148,765 purchased with capital of only $22,975.

The fifth and final subtlety with respect to net returns is that financing costs should most likely be considered even without any explicit repo borrowing. Any money used to purchase a bond had to have been raised at some cost (e.g., a bank paying money to attract deposits; a life insurance company compensating investors for the savings portions of their policies; a hedge fund offering a return on its assets under management). Furthermore, any money invested in a particular bond has an opportunity cost in the sense that this money might have been used to fund a different investment.

3.2 YIELD TO MATURITY

Chapter 2 made the point that rates are often more intuitive than prices in describing bond valuation. But, along the lines of that chapter, to describe the pricing of a 10-year bond in terms of semiannually compounded spot rates or forward rates requires 20 of those spot or forward rates. It is hardly surprising, therefore, that market participants prefer to quote the price of a bond and to think about its valuation with a single rate, namely, its yield to maturity.

The yield to maturity of a bond is the single rate such that discounting the bond's cash flows by that rate gives the bond's market price. Table 1.4 reported the price of the US Treasury 7.625s of 11/15/2022 for settlement in mid-May 2021 as 111.3969. Recalling that US Treasury bonds pay coupons semiannually so that this bond had three remaining payment dates, and noting the Treasury bonds are quoted with semiannually compounded yields, the yield to maturity of this bond, y, is defined as,[2]

$$111.3969 = \frac{3.8125}{\left(1+\frac{y}{2}\right)} + \frac{3.8125}{\left(1+\frac{y}{2}\right)^2} + \frac{103.8125}{\left(1+\frac{y}{2}\right)^3} \tag{3.5}$$

Equation (3.5) can be solved for y by some numerical method or with a financial calculator, giving a result of 0.0252%. Hence, trades of this bond could be quoted and consummated either at a price of 111.3969 or at a yield of 0.0252%.

More generally, let y denote the yield of a bond; let c denote its annual, dollar coupon payment; let T denote its maturity in years, which means

[2]This is only approximately correct because the settlement date is actually May 17, 2021. Hence, there is slightly less than one semiannual period to the first coupon payment date. Appendix A3.1 defines yield in this more general case.

there are $2T$ semiannual payments remaining; and let P denote its price, per 100 face amount, for settlement on a coupon payment date. To illustrate notation, note that for the 7.625s of 11/15/2022 settling in mid-May 2021, $c = 7.625$, $T = 1.5$, $P = 111.3969$, and $y = 0.0252\%$. Returning to the general case, then, P is given by,

$$P = \frac{\frac{1}{2}c}{\left(1+\frac{y}{2}\right)} + \frac{\frac{1}{2}c}{\left(1+\frac{y}{2}\right)^2} + \cdots + \frac{100+\frac{1}{2}c}{\left(1+\frac{y}{2}\right)^{2T}} \tag{3.6}$$

$$P = \frac{c}{2}\sum_{t=1}^{2T} \frac{1}{\left(1+\frac{y}{2}\right)^t} + \frac{100}{\left(1+\frac{y}{2}\right)^{2T}} \tag{3.7}$$

$$P = \frac{c}{y}\left(1 - \frac{1}{\left(1+\frac{y}{2}\right)^{2T}}\right) + \frac{100}{\left(1+\frac{y}{2}\right)^{2T}} \tag{3.8}$$

where Equation (3.7) just re-expresses (3.6) with summation notation, and Equation (3.8) follows from (3.7) and Equation (A2.16) of Appendix A2.2.

Equation (3.8) reveals three immediate implications about the price–yield relationship. First, when $c = 100y$, $P = 100$: when the coupon payment is the face amount times the yield, or, more simply, when the coupon rate equals the yield, the bonds sells for its face amount, or par. Second, when $c > 100y$, $P > 100$: when the coupon rate exceeds the yield, the bond sells at a *premium* to par. Third, when $c < 100y$, $P < 100$: when the yield exceeds the coupon rate, the bond sells at a *discount* to par.

Figure 3.1 illustrates the relationship between price and yield as described by Equations (3.6) to (3.8). Every point along the curves in the figure represents a bond with a particular coupon and maturity, and all bonds in the figure are priced at a yield of 1.5%. The thick black line, then, shows that bonds with a coupon rate of 1.5%, of any maturity, have a price of 100. Investors are willing to pay 100 for a bond that pays the going market coupon rate over time and then returns par at maturity.

Bonds with a coupon rate greater than the fixed yield of 1.5%, represented by the 3% and 6% gray, dashed lines in the figure, sell at a premium to par. With 30 years to maturity, the price a 6% bond yielding 1.5% is about 208, while the price of a 3% bond at that yield is about 136. Investors are willing to pay much more than par for these bonds because they pay above-market coupon rates for many years. Bonds with the same coupon rates but shorter maturities still command premiums, but smaller ones. For example, the prices of the 15-year bonds in the figure with 6% and 3% coupon rates are about 160 and 120, respectively. And, as maturities get very short, prices approach 100: while the 6% and 3% bonds do

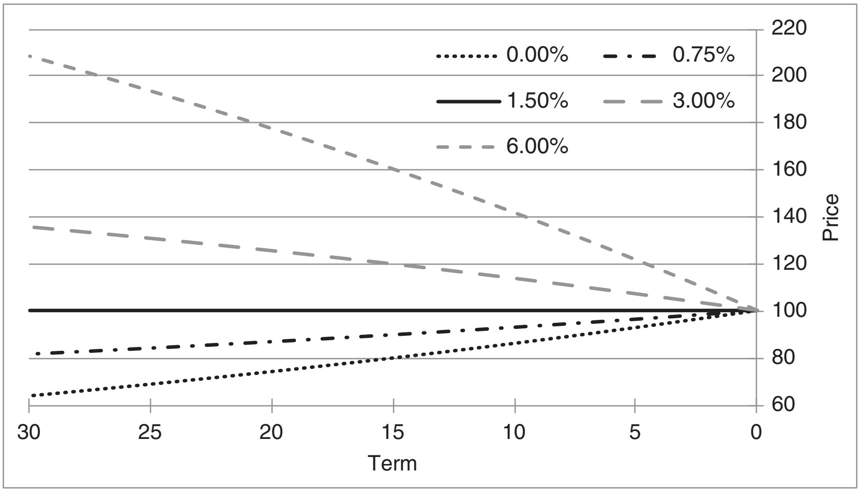

FIGURE 3.1 Prices of Bonds with Different Coupons and Maturities. All Yields Equal 1.5%.

pay above-market coupon rates, they do so for such a short time that their prices are not much above 100.

Following similar reasoning, the lines in Figure 3.1 representing bonds with coupon rates of 0% and 0.75% show prices all below par. The discount is greatest for the longest maturity bonds, for which investors receive below-market coupon rates for the longest times, with the prices of 30-year 0% and 0.75% bonds being about 64 and 82, respectively. Price is less discounted for shorter maturity bonds, to about 80 and 90 at 15 years, respectively, until bonds very near maturity sell for just under par. The trend of premium and discount bond prices to approach par as they mature is known as the *pull to par*. This behavior, of course, is only the trend, at a fixed yield of 1.5%. The actual price paths of these bonds as they mature are determined by their realized market yields.

3.3 YIELD AND RETURN

As mentioned earlier, yield to maturity is a convenient way to express price in terms of a single rate of interest. And the definition of yield as the single rate that equates the present value of a bond's cash flows to its price means that yield is a bond's *internal rate of return*. But how does yield relate to a bond's realized return?

As it turns out, yield is only weakly related to realized returns. Appendix A3.2 shows that a bond's *ex-post* return is equal to its initial yield if i) all of the coupons are reinvested at the initial yield, and ii) if the yield at the end of the investment horizon is the same as the initial yield. These very restrictive conditions significantly weaken the interpretation of yield as a predictor of holding period return.

To demonstrate this point with a concrete example, consider the return on the 7.625s of 11/15/2022 from mid-May 2021 to mid-May 2022. Begin with Equation (3.5), which gives that bond's yield as of the start of that period, and multiply both sides by $(1+y/2)^2$,

$$111.3969\left(1+\frac{y}{2}\right)^2 = 3.8125\left(1+\frac{y}{2}\right) + 3.8125 + \frac{103.8125}{\left(1+\frac{y}{2}\right)} \tag{3.9}$$

The left-hand side is the proceeds from an investment of the price of the bond, 111.3969, for one year – or two semiannual periods – at the semi-annually compounded rate y. The right-hand side is the value of the bond position at the end of the one-year horizon, which is the sum of three parts: the proceeds from the first coupon, received on November 15, 2021, and invested for six months to May 15, 2022; the coupon received on May 15, 2022; and the price of the bond on May 15, 2022, if its yield is y. In words then, Equation (3.9) says that the one-year return on the initial investment will be y if the first coupon is reinvested at y and if the bond, at the end of the year, is priced at a yield of y. If y in every term of (3.9) is the initial yield of the bond, that is, 0.0252%, then the bond earns that initial yield over the year. But if the coupon reinvestment rate or the yield of the bond at the end of year have changed over time, and are not equal to 0.0252%, then the bond likely does not earn that rate over the one-year horizon.

To see that a bond may not earn its initial yield even if held to maturity, multiply both sides of Equation (3.5) by $(1+y/2)^3$ to get,

$$111.3969\left(1+\frac{y}{2}\right)^3 = 3.8125\left(1+\frac{y}{2}\right)^2 + 3.8125\left(1+\frac{y}{2}\right) + 103.8125 \tag{3.10}$$

Now, the left-hand side is the proceeds from an investment of 111.3969 for 1.5 years. The right-hand side is the value of the bond position at maturity, that is, November 15, 2022, which is, again, the sum of reinvested coupons and the final coupon and principal payments. If y in every term of (3.10) is the initial yield, then the bond earns that initial yield to maturity. But if coupons are invested at a different rate, the bond does not earn its initial yield over its life.

The 7.625s of 11/15/2022 are useful in making the point that yield does not coincide with horizon return, but its short maturity, along with its low yield, may give a misleading impression of orders of magnitude. Consider, therefore, buying the US Treasury 2.375s of 5/15/2051 in mid-May 2021, at a price of 100.6875 or yield of 2.343%, and holding it to maturity. Along the lines of the previous paragraphs, if all coupons are reinvested at 2.343%, then the return on the initial investment over the 30 years is 2.343% per year. But if rates were to fall suddenly and remain low, so that all coupons are invested at 0%, or if the investor holds all coupon payments in a non-interest-bearing account, the return on the bond to maturity falls to 1.778% per year. If, on the other hand, rates were to rise suddenly and remain high, so that all coupons were reinvested at 5%, then the bond return rises to 3.207% per year. Reproducing these results is left as an exercise to the reader.

3.4 YIELD AND RELATIVE VALUE

This section argues that yield is not a reliable measure of relative value. In fact, if two bonds have the same maturity but different yields, it is not necessarily true that the bond with a higher yield is a superior investment. To explain this, the discussion turns to the *coupon effect*, starting with a simple numerical example and finishing with an empirical demonstration from the US Treasury bond market.

Say that the one-year spot rate is 0% and the two-year spot rate is 10%. Using the analytics of Chapter 2 and of this chapter, Table 3.1 gives the prices and yields of three two-year bonds with annual coupons of 0%, 5%, and 9.5023%.

For example, for the 5% bond,

$$91.7769 = \frac{5}{(1+0\%)} + \frac{105}{(1+10\%)^2} \tag{3.11}$$

$$= \frac{5}{(1+9.7203\%)} + \frac{105}{(1+9.7203\%)^2} \tag{3.12}$$

where Equation (3.11) uses the assumed spot rates to discount cash flows and Equation (3.12) follows from the definition of yield to maturity.

According to Table 3.1, the yield of the zero coupon bond is 10%. Because this bond makes only one payment at maturity, its market price can be found by discounting at either the two-year spot rate or at the bond's yield. Hence, those two must be equal at 10%.

TABLE 3.1 Prices and Yields of Two-Year Bonds When the One- and Two-Year Spot Rates are 0% and 10%, Respectively. Coupons and Yields Are in Percent.

Coupon	Price	Yield
0.0000	82.6446	10.0000
5.0000	91.7769	9.7203
9.5023	100.0000	9.5023

The yield of the 5% bond, however, is 9.7203%. Equations (3.11) and (3.12) show that the yield of 9.7203% has to summarize the term structure of spot rates, or, in other words, that discounting at the yield has to have the same effect as discounting the first cash flow by 0% and the second by 10%. The result of 9.7203% is between the two spot rates, but a lot closer to 10%: most of the bond's cash flow comes from the second payment, which includes principal, and that second payment is discounted at 10%.

Table 3.1 also reports that the 9.5023% bond, which sells for par, has a yield of 9.5023%. This yield is also between the spot rates of 0% and 10%, and is also a lot closer to 10% but not as close as for the 5% bond. Because the 9.5023% bond pays out relatively more of its value in the first cash flow, its yield is closer to 0% than is the yield of the 5% bond.

In general, the coupon effect can be summarized as follows: when the term structure of spot rates is upward sloping, bonds with higher coupons, which have more of their value discounted at shorter-term, lower spot rates, have lower yields to maturity. While not part of the discussion here, it is also true that when the term structure of spot rates is downward sloping, bonds with higher coupons have higher yields.

The coupon effect makes it very clear that spot rates are meant to describe pricing in a bond market as a whole, while yields are meant to describe the pricing of individual bonds. Furthermore, even though all of the bonds in the table are fairly priced relative to the term structure of spot rates, each bond has a different yield. Put another way, the fact that the 0% bond has the highest yield does not mean that it is the best investment. The fact that the par bond has the lowest yield does not mean that it is the worst investment.

Figures 3.2 and 3.3 show the coupon effect at work in the US Treasury market as of mid-May 2021. Figure 3.2 graphs the yields of all Treasury coupon bonds against their maturity dates. Some bond yields are clearly off the curve. The four points above the curve with maturities from 2040 to 2041 are relatively newly issued 20-year bonds. The Treasury stopped issuing new 20-year bonds in 1986 but started to do so again in May 2020. These four bonds, therefore, have coupons that reflect the current, low-rate environment. The other outstanding bonds with similar maturities were originally issued about 10 years ago as 30-year bonds, and, therefore, have relatively high coupons, which reflect the interest rate environment at the time they were issued.

Table 3.2 shows the coupons and yields of these four newly issued 20-year bonds, each paired with the coupons and yields of old 30-year bonds of the same maturity. As expected, because of the coupon effect and the upward-sloping term structure of rates, the high-coupon, old 30-year bonds have lower yields than the low-coupon, newly issued 20-year bonds. More analysis is needed, of course, to say that any one bond in this table is cheap or rich relative to another. But, because of the coupon effect, the fact

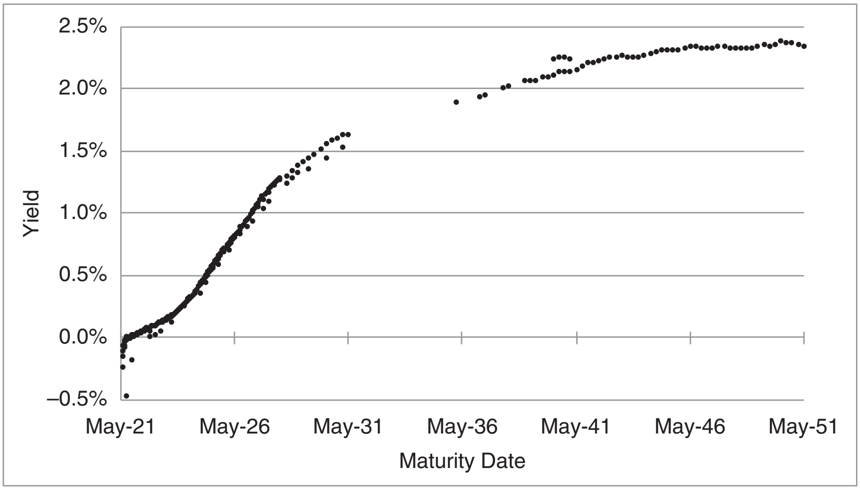

FIGURE 3.2 Yields of US Treasury Bonds, as of May 14, 2021.

TABLE 3.2 Yields of Selected US Treasury Bonds Maturing Between May 15, 2040, and February 15, 2041, as of May 14, 2021. Coupons and Yields Are in Percent.

	New 20-Year Bonds		Old 30-Year Bonds	
Maturity	Coupon	Yield	Coupon	Yield
05/15/2040	1.125	2.237	4.375	2.107
08/15/2040	1.125	2.245	3.875	2.138
11/15/2040	1.375	2.245	4.250	2.140
02/15/2041	1.875	2.236	4.750	2.133

that the new 20-year bonds have higher yields does not necessarily imply that they are better investments.

Returning to Figure 3.2, there are several bonds that mature on or before May 2031 that have yields noticeably below the curve.[3] To examine these bonds, Figure 3.3 zooms in on maturities on or before May 2031. The dots represent bonds with coupon rates less than or equal to 5%, while the plus signs represent bonds with coupon rates greater or equal to 5.25%. Once again, as expected because of the coupon effect and the upward-sloping term structure, the bonds with the higher coupons have lower yields.

[3]There are no bonds with maturities between May 2031 and February 2036 because the Treasury stopped issuing new 30-year bonds after May 2001 and started issuing them again in February 2006.

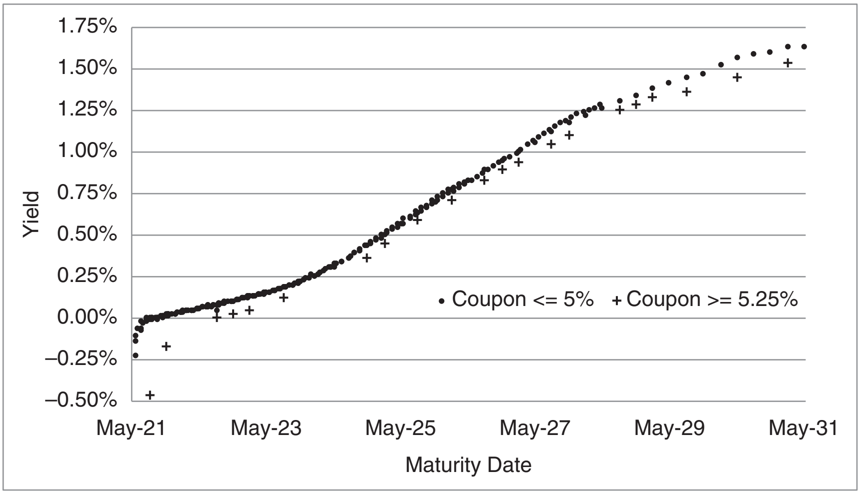

FIGURE 3.3 Yields of US Treasury Bonds Maturing in Less than 10 Years, as of May 14, 2021.

3.5 SPREADS

There are many fixed income securities. Their prices all depend in some way on market-wide interest rates, but also on idiosyncratic factors, like credit risk, tax treatment, and unique supply and demand considerations. To make sense of pricing in this context, market participants often quote spreads of one instrument or set of instruments relative to others.

Consider the European sovereign debt market. Because the market perceives the credit characteristics of each country's bond issues differently, some countries typically pay higher interest rates on their bonds than other countries. These differences could be described by comparing individual prices and yield. As of May 2021, for example, the Italian government bond or *BTP (Buoni del Tesoro Poliennali)* 0.60s of 08/1/2031 (issued in February 2021) sold at a price of 95.502, or a yield of 1.066%. At the same time, the German government bond or *Bund* 0.00s of 02/15/2031 (issued in January 2021) sold at a price of 101.300, or a yield of −0.132%. Market participants often find it more intuitive, however, to describe the market by saying that the 10-year BTP-Bund spread – in this case just the difference between the two yields – is 1.198%, or about 120 *basis points*. Note that one basis point is defined as 0.01%, so that 120 basis points is 1.20%. In any case, spreads of government bonds in many countries are often quoted against a Bund benchmark because Germany is widely regarded as having the strongest credit in the region.

Yield spreads can also be convenient for trading bonds that are less liquid than benchmark bonds. For example, in August 2020, Johnson & Johnson (JNJ) issued a number of long-term bonds, including its 20-year

2.10s due 09/1/2040. The pricing of this new issue was quoted as 75 basis points above the on-the-run 30-year Treasury bond, the 1.25s of 05/15/2050. Quoting yield against a Treasury benchmark is not only intuitive, but also facilitates trading in fast-moving markets. A fixed yield spread over a liquid Treasury bond allows the offering price of the JNJ bond to change smoothly and predictably with market levels until the new issue is actually sold.

While yield spreads are particularly easy to calculate and use, they are actually difficult to interpret carefully for two reasons. First, in many situations there is no liquid, benchmark bond with exactly the same maturity as the less-liquid bond. Taking the differences in yields, therefore, mixes differences in term with other differences, like credit. In the 20-year JNJ issue, for example, the spread of 75 basis points includes not only the credit risk of JNJ relative to the US Treasury, but also the difference between 20- and 30-year yields. Second, spreads of yields, even of the same maturity, inherit the weakness of yield described earlier and expressed as the coupon effect. Put another way, the yield spread between bonds of the same maturity can be misleading if their coupons are different and their underlying term structures have different shapes.

Bond spreads are a more careful and more meaningful formulation of expressing price differences as a spread of rates.[4] To illustrate, consider again the US Treasury 7.625s of 11/15/2022. Table 1.3 derives discount factors from a set of newly issued, benchmark Treasury bonds, which did not include the 7.625s of 11/15/2022. Table 1.4 then shows that the 111.3969 market price of the 7.625s of 11/15/2022 is 11.72 cents rich relative to its 111.2797 present value, which is computed using the discount factors derived from the benchmark bonds. The point of a bond spread is to express this 11.72 cents of richness as a spread of rates. Bond spreads can be computed relative to par, spot, or forward rates, but this section works with forward rates. The forward rates implied from the discount factors in Table 1.3 are 0.0154%, 0.1008%, and 0.1833%, for terms of 0.5, 1.0, and 1.5 years, respectively. Therefore, the 111.2797 present value of the 7.625s of 11/15/2022 is given by,

$$
\begin{aligned}
111.2797 = {} & \frac{3.8125}{\left(1+\frac{0.0154\%}{2}\right)} + \frac{3.8125}{\left(1+\frac{0.0154\%}{2}\right)\left(1+\frac{0.1008\%}{2}\right)} \\
& + \frac{103.8125}{\left(1+\frac{0.0154\%}{2}\right)\left(1+\frac{0.1008\%}{2}\right)\left(1+\frac{0.1833\%}{2}\right)}
\end{aligned}
\tag{3.13}
$$

[4]In the case of a bond with no embedded options, bond spreads are the same as *option-adjusted spreads*, which are described in Chapter 7.

The bond spread with respect to forward rates is defined as the spread that, when added to each forward rate, sets the bond's present value equal to its market price. In this case, denoting the spread by s,

$$
\begin{aligned}
111.3969 = {} & \frac{3.8125}{\left(1+\frac{0.0154\%+s}{2}\right)} \\
& + \frac{3.8125}{\left(1+\frac{0.0154\%+s}{2}\right)\left(1+\frac{0.1008\%+s}{2}\right)} \\
& + \frac{103.8125}{\left(1+\frac{0.0154\%+s}{2}\right)\left(1+\frac{0.1008\%+s}{2}\right)\left(1+\frac{0.1833\%+s}{2}\right)}
\end{aligned}
\tag{3.14}
$$

Solving, $s = -0.0727\%$, or about minus seven basis points. The spread is negative because the bond is rich; that is, the market price is above the present value as computed with the benchmark curve. In other words, the benchmark forward curve has to be shifted down to recover the relatively high market price.[5]

A bond's spread can be interpreted as its extra return, relative to the benchmark bonds, if rates "stay the same," or if interest rate risk is hedged away. This interpretation is developed in the next section and in Chapter 7.

Because bond spread can be interpreted as an indicator of relative value, many practitioners routinely compute bond spreads, whether it be spreads of various European government bonds to German Bunds, spreads of corporate bonds to government bonds or swaps, spreads of government bonds to benchmark government bonds, or even spreads of government bonds to swaps. With respect to the last of these, it had once been standard to compute spreads of government bonds only to the most liquid government bonds. But as LIBOR swaps became more and more liquid, and as the idiosyncratic pricing of individual government bonds became more apparent, market participants began to compute spreads of government bonds to LIBOR swaps as well. SOFR swaps, which are now replacing LIBOR swaps, may take on the same role in the future.

---

[5]In some contexts, like the corporate bond market, practitioners often define a term structure of spreads; that is, unlike in Equation (3.14), they allow a different spread to be added to each forward rate. This could capture, for example, the long-term bonds of a particular corporation selling at a higher spread to Treasury bonds than its short-term bonds. The discussion continues here, however, with a single spread.

3.6 APPLICATION: SPREADS OF HIGH-COUPON TREASURIES

To illustrate the use of bond spreads in assessing relative value, the top panel of Table 3.3 and the gray lines in Figure 3.4 report the bond spreads of selected Treasury bonds relative to benchmark Treasury bonds and also relative to SOFR swaps. The benchmark Treasuries, by definition, have zero spreads relative to themselves. Consistent with the analysis of high-coupon bonds in terms of the law of one price in Chapter 1, high-coupon bonds are rich relative to benchmark Treasuries, with bond spreads ranging from −0.8 to −18.9 basis points. Traders and investors can use these spreads in their investment decisions. Because these bonds are rich, buying them sacrifices return relative to buying benchmark bonds, implying that there needs to be some reason, not considered here, to purchase these bonds. Conversely, because these bonds are rich, selling or shorting them picks up return relative to benchmark bonds, absent some factor not considered in this analysis.

As mentioned already, market practitioners sometimes use swaps as a benchmark against which to compute government bond spreads. For the bonds in this example, spreads to SOFR swaps are reported in the bottom

TABLE 3.3 Spreads of Selected US Treasury Bonds to Benchmark Treasuries and to SOFR Swaps, as of May 14, 2021. Coupons Are in Percent, and Spreads Are in Basis Points.

	Benchmark Treasuries		High-Coupon Treasuries	
Maturity	**Coupon**	**Spread**	**Coupon**	**Spread**
		Spreads to Benchmark Treasuries		
11/15/2021	2.875	0.0	8.000	−18.9
11/15/2022	1.625	0.0	7.625	−7.3
11/15/2024	2.250	0.0	7.500	−5.8
11/15/2026	2.000	0.0	6.500	−0.8
11/15/2027	2.250	0.0	6.125	−2.4
11/15/2028	3.125	0.0	5.250	−2.1
		Spreads to SOFR Swaps		
11/15/2021	2.875	−1.9	8.000	−20.8
11/15/2022	1.625	3.1	7.625	−4.2
11/15/2024	2.250	5.9	7.500	−0.2
11/15/2026	2.000	16.0	6.500	14.4
11/15/2027	2.250	22.6	6.125	19.0
11/15/2028	3.125	26.4	5.250	23.5

Source: Author Calculations.

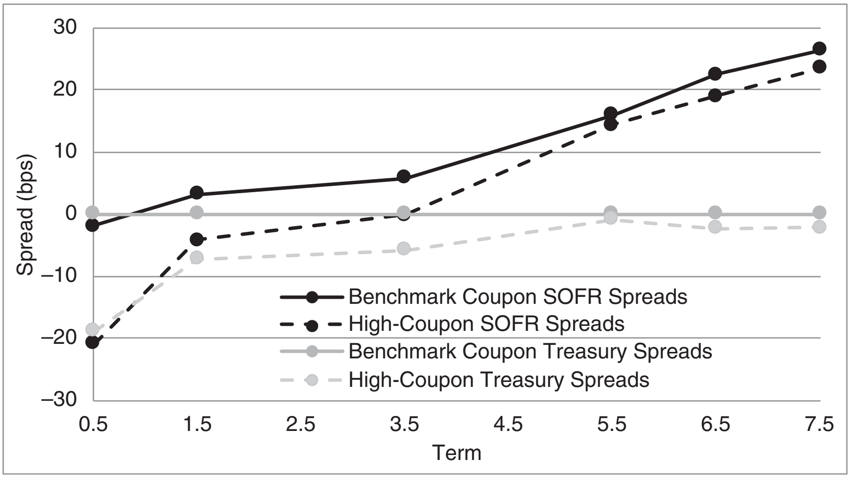

FIGURE 3.4 Spreads of Selected US Treasury Bonds to Benchmark Treasuries and to SOFR Swaps, as of May 14, 2021.

panel of Table 3.3 and with the black lines in Figure 3.4. With respect to the relative pricing of high-coupon bonds, the SOFR spreads tell the same story as the Treasury spreads: high-coupon bonds have lower (and more negative) spreads and, therefore, are rich to benchmark Treasury bonds. In fact, the differences between the spreads of high-coupon bonds and of benchmark bonds are about the same whether measured against benchmark bonds or SOFR swaps. With respect to relative pricing across term, however, the results are quite different. Measured against SOFR swaps, longer-term bonds – both high-coupon and benchmark – are much cheaper than shorter-term bonds. Consider, for example, the spreads of the 5.25s of 11/15/2028 and the 8s of 11/15/2021. The difference in their SOFR spreads is 23.5 – (–20.8) or 44.3 basis points, while the difference in their Treasury spreads is –2.1 – (–18.9) or only 16.8 basis points. And the difference in SOFR spreads between the 3.125s of 11/15/2028 and the 2.875s of 11/15/2021 is 26.4 – (–1.9) or 28.3 basis points, while the difference in their Treasury spreads, because they are both benchmark bonds, is, by definition, zero. Hence, longer-term bonds look much cheaper relative to short-term bonds when measured against SOFR than against Treasuries.

To understand this difference in relative valuations, Figure 3.5 graphs the two forward curves. Past a term of 2.5 years, the SOFR forward curve is not as steep as the Treasury forward curve. This means that the SOFR curve assigns relatively high prices to longer-term Treasury bonds and, consequently, makes their market prices look relatively cheap. Equivalently, SOFR spreads at longer terms need to be larger to get pricing back up to

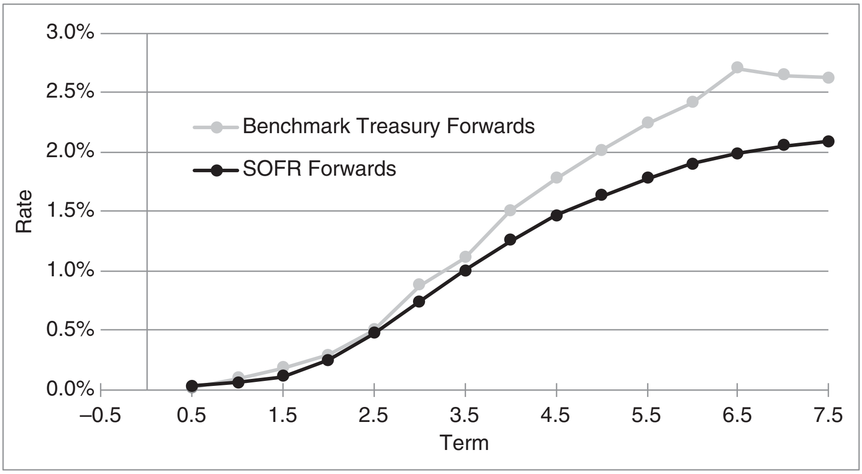

FIGURE 3.5 Benchmark Treasury and SOFR Forward Rate Curves, as of May 14, 2021.

the levels of Treasury forwards. The lesson here to practitioners is that using spreads against a curve from another market may have liquidity advantages, as mentioned earlier, but can complicate the interpretation of spreads and trades based on those spreads. In the present case, the relative cheapness of longer-term, high-coupon Treasuries, as indicated by their relatively large SOFR spreads, is actually a combination of their relative cheapness against the Treasury curve and the relative flatness of the SOFR curve. Furthermore, trades that attempt to lock in these spreads of Treasury bonds against SOFR spreads, called *asset swaps*, require trades not only in Treasuries but in matched-maturity SOFR swaps as well. Chapter 14 discusses asset swap trades in greater detail.

3.7 UNCHANGED RATE SCENARIOS FOR P&L ATTRIBUTION

This section and the next describe how *ex-post* P&L or returns can be broken down into components so as to understand the sources of profits or losses from a trade or investment. One of these components, *carry–roll-down*, is the change in value or the return of a bond or portfolio of bonds over time, assuming that “rates do not change” over the P&L or return horizon. The purpose of this section is to present two commonly assumed scenarios to represent rates not changing: *realized forwards* and an *unchanged term structure*.

TABLE 3.4 Realized Term Structure of Treasury Forward Rates on Various Dates, as of November 13, 2020. Rates Are in Percent.

Term	11/13/20	05/14/21	11/15/21	05/14/22
0.5	0.1013	0.1746	0.2429	0.2185
1.0	0.1746	0.2429	0.2185	
1.5	0.2429	0.2185		
2.0	0.2185			

Table 3.4 illustrates realized forward scenarios as of November 13, 2020, with sequential six-month horizons. The 11/13/2020 column gives the then-prevailing US Treasury term structure of six-month forward rates. Using the notation of Chapter 2, the first six-month forward rate, $f(0.5)$, is 0.1013%; the six-month rate, six months forward, $f(1)$, is 0.1746%; the six-month rate, one year forward, $f(1.5)$, is 0.2429%; etc. The next three columns give the assumed term structures of forward rates on subsequent dates. Note that the 11/13/2020 date is selected as a business day approximately six months before 05/14/2021, which is the pricing date for many of the examples in this text. The 11/15/2021 and 05/14/2022 dates are somewhat arbitrarily chosen as business days approximately six months apart.

Realized forward scenarios assume that today's six-month rate, t years forward, will be the six-month rate t years from now. In terms of the table, the November 13, 2020, six-month rate, six months forward, 0.1746%, is assumed to be the six-month rate in six months, that is, on May 14, 2021. The November 13, 2020, six-month rate, one year forward, 0.2429%, is assumed to be the six-month rate on November 15, 2021; and the November 13, 2020, six-month rate, 1.5 years forward, 0.2185%, is assumed to be the six-month rate on May 14, 2022. Furthermore, all along the way, the forward rates slide down the curve. The November 13, 2020, six-month rate, one-year forward, 0.2429%, becomes the six-month rate, six months forward, on May 14, 2021, before becoming the six-month rate on November 15, 2021. And the November 13, 2020, six-month rate, 1.5 years forward, 0.2185%, becomes the six-month rate, one year forward, on May 14, 2021, and the six-month rate, six months forward, on November 15, 2021, before becoming the six-month rate on May 14, 2022.

The scenario of an unchanged term structure assumes that all forward rates remain the same period after period. In terms of Table 3.4, this scenario assumes that the term structure of forward rates on November 13, 2020, is the same as the term structure on May 14, 2021, November 15, 2021, and May 14, 2022. In terms of rates, $f(0.5)$ stays at 0.1013%; $f(1)$ stays at 0.1746%; and so forth.

TABLE 3.5 Return on a T-Year Zero Coupon Bond Under the Scenarios of Realized Forward Rates and an Unchanged Term Structure, with Annual Periods.

	Realized Forwards	Unchanged Term Structure
P_0	$\frac{1}{(1+f(1))(1+f(2))\cdots(1+f(T-1))(1+f(T))}$	$\frac{1}{(1+f(1))(1+f(2))\cdots(1+f(T-1))(1+f(T))}$
P_1	$\frac{1}{(1+f(2))\cdots(1+f(T-1))(1+f(T))}$	$\frac{1}{(1+f(1))(1+f(2))\cdots(1+f(T-1))}$
$\frac{P_1}{P_0}-1$	$f(1)$	$f(T)$

Chapter 8 characterizes the realized forward scenario as the *pure expectations hypothesis* and the unchanged term structure scenarios as the *pure risk premium hypothesis*. For now, however, Table 3.5 gives some insight into bond returns under the two scenarios by deriving the annual return on a zero coupon bond of maturity T, with annual periods. The price of the bonds today, P_0, are both given by discounting their maturing face value of one unit of currency by the annual forward rates $f(1)$ through $f(T)$. In one year, the bonds are both $(T-1)$-year zeros, and their prices, P_1, differ according to the assumed scenario. Under realized forwards, the first forward rate becomes $f(2)$, etc., out to $f(T)$, giving the P_1 value in that column of the table. Under an unchanged term structure, the appropriate forward rates are $f(1)$ through $f(T-1)$: $f(T)$ drops out, because the bond matures in $T-1$ years. The last row of the table gives the one-year return of the bond, which is simply $(P_1 - P_0)/P_0$, or $P_1/P_0 - 1$.

The table shows that, under realized forwards, the one-year return on the T-year zero is the one-year rate, $f(1)$. Appendix A3.3 shows that this result is very general: under realized forwards, the return on any bond equals the short-term rate. And if the bond trades at a constant spread to the benchmark curve, the return equals the short-term rate plus that spread. To the extent the scenario is considered reasonable, this result is useful for P&L attribution, because any return different from the short-term rate is attributable not to the passage of time, but to other factors, like changes in rates or spreads.

Under the assumption of an unchanged term structure, the one-year return on the T-year zero is the initial forward rate from $T-1$ to T. This result does not generalize as easily as the return under realized forwards, and the intuition behind both results becomes clear in Chapter 8. For the purpose of this section, however, suffice it to say that if a trader or investment manager finds the unchanged term structure scenario a more reasonable expression of “rates stay the same,” then returns under that scenario can be considered a baseline, and return deviations from that baseline are attributed to other factors.

3.8 P&L ATTRIBUTION

As mentioned earlier, breaking down P&L or return into component parts is extremely useful for understanding how money is being made or lost in a trading book or investment portfolio. In addition, attribution reports can often catch errors of many sorts (e.g., buying or selling the wrong bond; buying instead of selling or vice versa; incorrect market or position data feeds).

P&L attribution can be quite detailed, but, for the purposes of this chapter, the focus is on a single bond rather than a portfolio; the holding period is fixed at a coupon payment period; and attribution is to four factors: *cash carry*, *carry–roll-down*, changes in rates, and changes in spreads.[6]

Cash carry captures interim cash flows from a bond investment, typically from coupons and from financing arrangements (i.e., borrowing funds to buy bonds or borrowing bonds to short them). The example of this section ignores financing, however, which is the subject of Chapter 10.

Carry–roll-down captures the P&L or return from changes in price when neither rates nor spreads change. The previous section described how rates not changing might be modeled as realized forwards or as an unchanged term structure. The example of this section assumes realized forwards.

Because "carry" and "roll-down" are used differently by different practitioners, the nomenclature adopted here requires some elaboration. Broadly speaking, carry is meant to convey how much a position earns due to the passage of time, holding everything else constant. A clean example is a par bond when the term structure of interest rates is flat and unchanging. In that case, because the bond's price is always par, its carry is clearly its coupon income minus costs of financing. Another clean example is a premium bond when the term structure, again, is flat and unchanging. In that case, because the bond's price is pulled to par over time (Figure 3.1), its carry is clearly its coupon income minus the decline in price minus costs of financing.

Roll-down, broadly speaking, is meant to convey how much a position earns as it matures and is valued at a different part of the term structure, holding everything else constant. A clean example of this concept is a forward loan. Referring to the rates in Table 3.4, say that, on November

---

[6]More detailed analyses, instead of considering rates changes as a single change from one term structure to the next, consider, for instance, changes to the level of rates and the steepness of the term structure separately, or consider changes to individual segments of the term structure separately. These choices can be understood more fully in the context of the multi-factor risk metrics in Chapters 5 and 6. More detailed analyses can also include more spreads, for example, the spread of Treasury rates to repo or financing rates; the spread of swap rates to Treasury rates; and the spread of various corporate rates to swap rates.

13, 2020, an investor makes a six-month loan, one year forward, at 0.2429%, and has adopted the unchanged term structure scenario for P&L attribution. Six months later, the investor's position will have matured to a six-month loan, six months forward to be valued at the "unchanged" rate of 0.1746%, which is the six-month rate, six months forward, on November 13, 2020. Hence, holding everything else constant, the forward loan has gained in value: the investor had locked in a loan rate of 0.2429% on November 13, 2020, and, six months later, the loan is valued at 0.1746%. In other words, the loan increased in value because it rolled down the curve.

While there are some cases in which it is clear how to separate carry from roll-down, there are many cases in which it is not so clear. Consider a premium bond when the term structure is upward sloping and unchanging. The resulting P&L would be a combination of carry – coupon, pull-to-par, and financing costs – and roll-down, as the bond's cash flows are discounted at lower rates.

This text takes the position that the important division in P&L attribution is between what happens due solely to the passage of time and what happens when rates and spreads change. Furthermore, to be consistent with the nomenclature in forwards and futures markets, discussed in Chapter 11, the phrase cash carry is preserved to mean coupon income minus financing costs. Therefore, carry–roll-down here denotes P&L or return due to the passage of time excluding cash carry. The name reflects carry–roll-down's being a mix of what might otherwise be classified as either carry or roll-down.

The remaining two components of P&L or return described in this chapter, those due to changes in rates and to changes in spreads, are relatively self-explanatory and are discussed presently.

The text now turns to a detailed P&L attribution for the US Treasury 7.625s of 11/15/2022 over the period November 13, 2020, to May 14, 2021. Note that both of these pricing dates are Fridays, so that settlement is on November 16, 2020, and May 17, 2021, respectively. Therefore, full prices given here include one and two days of accrued interest, respectively, and an investor over this holding period does not receive the coupon paid on November 15, 2020, but does receive the coupon paid on May 15, 2021.

The computation and results of the P&L attribution are given in Tables 3.6, 3.7, and 3.8. The first row of Table 3.6 gives the market price and spread of the 7.625s of 11/15/2022 as of the beginning of the period, November 13, 2020. This spread of −1.16 basis points is relative to the benchmark Treasury term structure on November 13, 2020, which is reported in the second column of Table 3.7. The computation of spread from market price is not shown here but can be computed along the same lines as Equation (3.14).

The second row of Table 3.6 computes a price for the bond as of the end of the period, May 14, 2021, assuming realized forwards – given in the third column of Table 3.7 – and an unchanged spread of −1.16 basis points.

TABLE 3.6 Data for the P&L Attribution of the 7.625s of 11/15/2022 from November 13, 2020, to May 14, 2021. Spreads Are in Basis Points, and Percent Change Is Relative to Price on November 13, 2020.

Pricing Date	Term Structure	Spread	Price	Change	Percent Change
11/13/2020	11/13/2020	−1.16	114.87654		
5/14/2021	Realized Forwards	−1.16	111.11555	−3.76099	−3.2739%
5/14/2021	5/14/2021	−1.16	111.29847	0.18292	0.1592%
5/14/2021	5/14/2021	−7.27	111.39690	0.09843	0.0857%

TABLE 3.7 Term Structures of Forward Rates for the P&L Attribution of the 7.625s of 11/15/2022 from November 13, 2020, to May 14, 2021. Rates Are in Percent.

	Term Structures of Forward Rates		
Term	11/14/20	Realized Forwards	5/14/21
0.5	0.1013	0.1746	0.0154
1.0	0.1746	0.2429	0.1008
1.5	0.2429	0.2185	0.1833
2.0	0.2185		

Once again, this spread calculation can be done along the lines of Equation (3.14). The change in price from 114.88, in the first row of Table 3.6, to 111.12, in the second row of the table, represents what happens to the bond price due to the passage of time alone: the realized forwards scenario is taken here to mean no change in rates, and the spread is also unchanged. Therefore, the dollar difference in these prices, −3.76, and the difference as a percentage of the initial price, −3.27%, is attributed to carry–roll-down and reported as such in Table 3.8. The carry–roll-down is large and negative because the bond has a coupon rate very much above the market rate and, consistent with the discussion in Section 3.2, its price is pulled down to par over time.

The third row of Table 3.6 computes a price for the bond as of May 14, 2021, assuming the actual Treasury term structure on that date, given in the fourth column of Table 3.7, and a still-unchanged spread from the beginning of the period. Yet again, this price can be computed along the lines of Equation (3.14). The difference between the resulting price of 111.30 and the price of 111.12 in the second row represents the P&L due to a change in rates – from the realized forwards scenario of "no change" to the actual term structure on May 14, 2021. The differences of 0.18 and 0.16% are reported in Table 3.8 as the rates component row of the attribution. This component

TABLE 3.8 P&L Attribution of the 7.625s of 11/15/2022 from November 13, 2020, to May 14, 2021. Returns Are in Percent.

Component	P&L	Return
Cash Carry	3.81250	3.3188
Carry–Roll-Down	−3.76099	−3.2739
Rates	0.18292	0.1592
Spread	0.09843	0.0857
Total	0.33286	0.2898

of the attribution is positive because, as is clear from Table 3.7, rates fell from the realized forwards scenario to the actual rates on May 14, 2021.

The fourth and final row of Table 3.6 gives the market price of the bond on May 14, 2021, along with its −7.27 basis-point spread relative to the benchmark Treasury term structure on that date. This spread is actually computed in Equation (3.14). The difference between the May 14, 2021, market price and the price in the third row of the table represents the spread component of the attribution, as the only change from the third to the fourth rows is the spread. The differences of 0.10 and 0.09% are reported as such in Table 3.8. This component is positive because the bond's spread fell from November 13, 2020, to May 14, 2021; that is, the bond became richer relative to the benchmark Treasury curve.

Table 3.8 summarizes the P&L and return attribution of the 7.625s of 11/15/2022 over the six-month period from November 13, 2020, to May 14, 2021. The first row, which has not yet been discussed, is the cash carry of 3.8125, that is, the coupon payment on May 15, 2021. (Recall that because the pricing date is on May 14 and the settlement date on May 17, the investor receives the May 15 coupon payment.) As a whole, the table breaks down the total dollar P&L of 0.33 and the total return 0.2898% into the component parts of cash carry, carry–roll-down, rates, and spread. Note that the sum of the cash carry and carry–roll-down return components equals 0.0449%, which is the six-month rate plus the bond spread on May 14, 2021, all divided by two, that is (0.1013% − 0.0116%)/2. This is, of course, no coincidence. As highlighted in the previous section, the annual return of any bond under the realized forward scenario is the short-term rate plus the spread, or half of that over a six-month horizon.

A trader or investor can use P&L and return attribution in several ways. First, as mentioned earlier, any surprises here might be due to errors in building positions, booking positions, or data feeds. Second, because cash carry and carry–roll-down can be computed at the beginning of the period, the trader or investor can decide whether the cash flow properties of the position are acceptable. More specifically, a trade or investment that is expected to do well in the long run but expected to bleed cash in the short run may or may not be acceptable in all institutional settings. In the present example, it is known at the beginning of the period that the "no change" case will result in a positive cash flow of 3.81250 – 3.76099 or about 0.05 cents before financing costs. Once financing costs are considered, this central case might be negative and might, in that case, be unacceptable.

A third use of P&L and return attribution is as an *ex-post* check on performance. Did the investor or trader have an *ex-ante* view on rates and, if so, how did realized rate changes compare with that view? Was there a view about how the spread of the bond would evolve and, again, if so, how did that view compare with the realized spread? Many traders and investors like to ask the following, when setting up a position: What could happen to this position that would convince me that I don't understand the relevant market forces? They would then promise themselves to exit the position if it turned out that their understanding was flawed. P&L and return attribution is a valuable tool in that discipline by revealing exactly where money is made or lost.

CHAPTER 4

DV01, Duration, and Convexity

This chapter, along with Chapters 5 and 6, are about measuring and hedging interest rate risk. Market participants need to understand how fixed income prices change when interest rates change to take a view on the future level or term structure of interest rates, to ensure that a portfolio of assets keeps pace with a portfolio of liabilities, or to hedge one fixed income instrument or portfolio with another. But how exactly should a "change" in rates be defined? Chapters 1 through 3 show that pricing in fixed income markets can be expressed in terms of discount factors, par rates, spot rates, forward rates, yields, and spreads. Which of these quantities should be assumed to change and by how much?

In answering this question, there are two overarching trade-offs: simplicity versus empirical reality, and robustness versus model dependence. With respect to the first trade-off, a market maker who is hedging a long position in a 9.75-year bond with a short position in a 10-year bond over a short time period might reasonably rely on the simplest of frameworks, namely, that the yields of the two bonds will move up or down by the same number of basis points, that is, in *parallel*. A swap desk, by contrast, which is managing the risk of a portfolio of long and short swap positions of different terms, has to account for the reality that changes in rates across the term structure do not move perfectly in sync. With respect to robustness versus model dependence, a pension fund, which hedges the present value of its liabilities relative to the value of its assets, tends to prefer frameworks that are not very sensitive to assumptions about how rates of different terms vary relative to one another. An actively managed fixed income mutual fund, exchange-traded fund, or hedge fund, by contrast, which is in the business of taking views on the level of rates and on the shape of the term structure, might gravitate toward frameworks designed to incorporate subjective views.

The *one-factor* metrics and hedges described in this chapter are relatively simple along the lines just discussed, because they assume that rates of all terms move up and down together in some fixed relation. For example, if the 30-year par rate changes by one basis point, it might be assumed that the 10-year par rate changes by 0.99 basis points, in the same direction, while the five-year rate changes by 0.8 basis points, in the same direction. The parallel-shift assumption for bond yields is another simple framework described in this chapter.

The first several sections of this chapter introduce the metrics most often used to describe interest rate risk, namely, *DV01*, *duration*, and *convexity*, and illustrate these metrics by showing how to hedge Norfolk Southern Company's 100-year or *century* bond with a shorter-maturity Treasury bond. The next sections specialize the discussion to yield-based versions of these metrics, which, while more restrictive in terms of their underlying assumptions, are very intuitive and widely used in practice. The final section presents a stylized case of asset–liability management at a pension fund, which requires choosing between *barbell* and *bullet* asset portfolios. Chapters 5 and 6 continue the discussion of risk metrics and hedges with *multi-factor* and explicitly empirical approaches, respectively.

4.1 PRICE–RATE CURVES

The three bonds in Table 4.1 are used in this and subsequent sections to illustrate concepts. First, the Norfolk Southern Company (NSC) issued $600 million of the 4.10s of 05/15/2121 in May 2021 with a maturity of 100 years. Sales of *century* bonds are rare, but, with rates at historically low levels, bonds with very long maturities are being issued more often. The second and third bonds are the US Treasury 1.625s of 05/15/2026, issued in May 2016 with a bit less than $70 billion outstanding, and the 1.625s of 11/15/2050, issued in May 2020 with about $86 billion outstanding. The yields of these three bonds as of mid-May 2021 are given in the table, along with their spreads to the High-Quality Market-Weighted (HQM) corporate bond yield curve. The HQM curve, published by the US Treasury, is designed to be representative of corporate bonds rated A or above and is intended for use by pension funds in computing the present value of their liabilities.

As discussed in the introduction, a "change" in rates has to be defined in order to compute interest rate risk metrics. For current purposes, the following assumptions are made. First, the base curve is the HQM par rate

TABLE 4.1 Selected Bonds with Indicative Levels as of Mid-May 2021. Spreads Are Versus Par Rates from the High-Quality Market-Weighted Corporate Bond Curve, as of May 2021. Coupons and Yields Are in Percent. Spreads Are in Basis Points.

Issuer	Coupon	Maturity	Yield	Spread
Norfolk Southern Company	4.10	05/15/2121	4.103	66
US Treasury	1.625	05/15/2026	0.823	−41
US Treasury	1.625	11/15/2050	2.363	−97

curve. Second, the spreads of the NSC and US Treasury bonds relative to the HQM curve, as listed in Table 4.1, do not change as the HQM curve changes. Third, the change in the HQM par rate of any term is a fixed proportion of the change in the 30-year par rate. For example, if the 30-year par rate moves up by one basis point, the 10-year rate moves up by 0.99 basis points; the five-year rate by 0.80 basis points; the two-year rate by 0.37 basis points. These proportions are meant to capture the broad empirical fact, discussed further in Chapter 6, that short-term rates are less volatile than long-term rates. Figure 4.1 graphs the assumed shifts to the HQM par curve, which, for visibility, are scaled to a change of plus or minus 50-basis points in the 30-year par rate. It can easily be seen that, as intended, short-term rates move by less than long-term rates, and that the shift is close to a parallel shift for terms greater than 10 years.

Equipped with these assumptions, the prices of the bonds in Table 4.1 can be computed after the 30-year HQM par rate changes by, say, one basis point. More specifically: i) given the one-basis-point change in the 30-year par rate, calculate the change in par rates of all other terms along the lines of the previous paragraph; ii) for each bond, add its spread to the shifted par rates; iii) for each bond, convert the shifted par plus spread curve into discount factors; and iv) reprice each bond.

Figure 4.2 shows the results of these calculations. The horizontal axis gives changes in the 30-year HQM par rate, and the vertical axis gives corresponding bond prices. Note that each rate change is assumed to be instantaneous; that is, no time passes as the rate changes from its current to its new level. In any case, as expected, prices increase as rates fall and

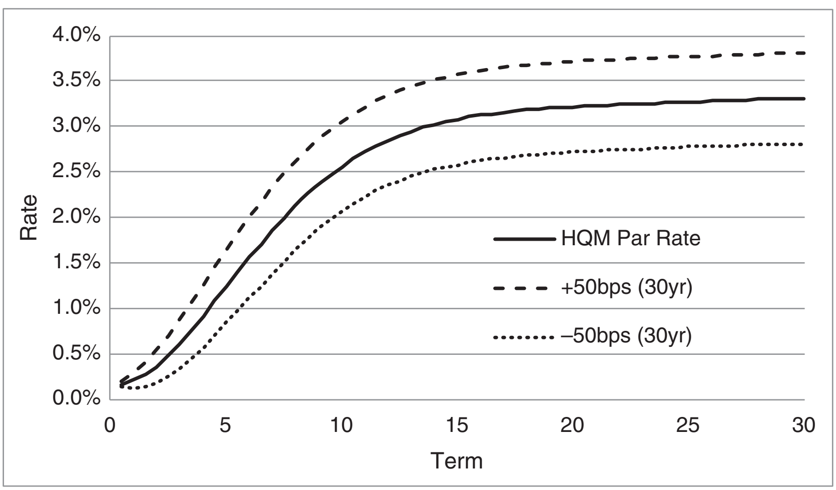

FIGURE 4.1 Sample Shifts to HQM Par Rates.

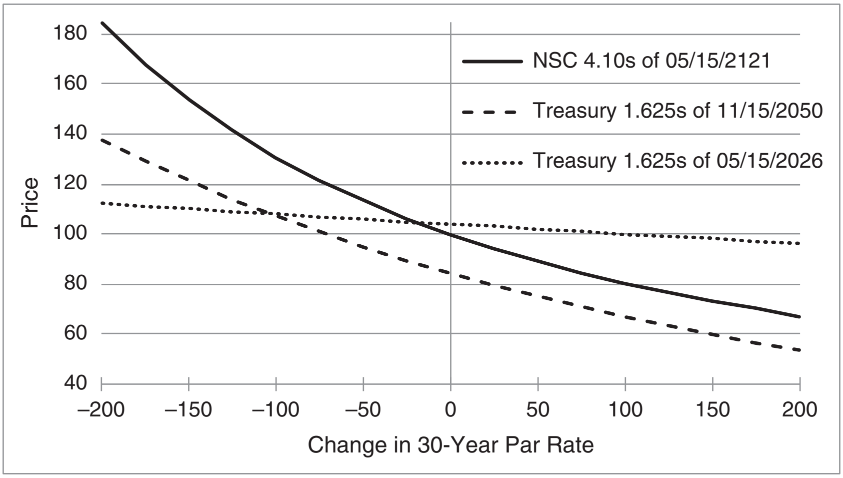

FIGURE 4.2 Price–Rate Curves for the Bonds in Table 4.1, as of Mid-May 2021.

decrease as rates rise. But the figure also shows that the three bonds do not have the same sensitivity to changes in rates. The price of the Treasury 1.625s of 05/15/2026, which has five years remaining to maturity, declines relatively mildly as rates increase; that is, its price is least sensitive to changes in rates. By the same token, the price of the Treasury 1.625s of 11/15/2050, which matures in 29.5 years, is more sensitive to rates, and the NSC century bond is the most sensitive of all. These differences in price sensitivities are quantified in the next sections as differences in DV01 or duration. Finally, the figure shows that the five-year Treasury price–rate curve is not far from a line, whereas the price–rate curve of the 29.5-year Treasury, and even more so of the NSC century bond, have noticeable curvature. These differences are quantified in subsequent sections as differences in convexity.

4.2 DV01

DV01 is an acronym for the *dollar value of an '01*, that is, the change in the price of a bond for a change in rates of 0.01%, or one basis point. As described in this section, DV01 can be applied in the context of any one-factor model of changes in rates across the term structure. Practitioners typically use the term DV01, however, to mean yield-based DV01, which is a narrower concept described in Section 4.7.

Table 4.2 illustrates the calculation of DV01 for the three bonds introduced in Table 4.1. Three prices are given for each bond. The price in column three is as of the pricing date in mid-May 2021. The prices in columns two and four are the prices after a change in the 30-year HQM par rate of

TABLE 4.2 Calculating DV01 for Bonds in Table 4.1, as of Mid-May 2021.

Bond	Price −1bp	Price	Price +1bp	Slope	DV01
NSC of May 2121	100.1801	99.9390	99.6990	−2,406	0.241
Treasury of May 2026	103.9621	103.9219	103.8817	−402	0.040
Treasury of Nov. 2050	84.5899	84.3906	84.1919	−1,990	0.199

−1 or +1 basis point, respectively. All other par rates fall or rise as well, of course, as discussed in the previous section. Given these prices, the first step in computing DV01 is to compute the slope of the price–rate curve, that is, the change in price divided by the change in rate. For the NSC 4.10s of 05/15/2021, the slope around the current market level is,

$$\frac{99.6990 - 100.1801}{0.02\%} = -2{,}406 \tag{4.1}$$

The left-hand side of (4.1) is the difference between the bond prices at +1 and −1 basis point divided by the change in rates of $+1 - (-1) = 2$ basis points, or 0.02%. Note that the current price – corresponding to a shift of 0 basis points – is not used in this calculation. From a numerical perspective, it is more accurate to estimate the slope of the curve at the current rate with shifts of +1 and −1 basis point, which are centered around the current rate, rather than with shifts of −1 and 0 basis points, which are centered slightly below the current rate, or with shifts of 0 and +1 basis point, which are centered slightly above the current rate.

While Equation (4.1) gives −2,406 as the slope of the price–rate curve, its scale – the change in price per unit change in rates, that is, a change of 1.0 or 100% or 10,000 basis points – is not very intuitive. Much more intuitive and useful is the change in price for a one-basis-point change in rates. To this end, the slope can be divided by 10,000 to give, rounded to three decimal places, −0.241. Furthermore, because the prices of almost all fixed income products fall as rates increase, it is conventional to drop the minus sign as understood. These two adjustments to the slope calculated in Equation (4.1) give the bond's DV01,

$$-\frac{1}{10{,}000}\frac{99.6990 - 100.1801}{0.02\%} = 0.241 \tag{4.2}$$

In words, a DV01 of 0.241 means that 100 face amount of the bond increases in price by 0.241 dollars or 24.1 cents when rates decrease by one basis point. Note that DV01 measures a price change per 100 face amount because the prices in Table 4.2 are per 100 face amount. While DV01, like price, is typically quoted per 100 face amount, it is occasionally useful to quote the DV01 of particular position. In these situations, DV01 is explicitly quoted in units of currency. For example, because the DV01 of the NSC century bond is 0.241 per 100 face amount, a position of $10 million has a DV01 of,

$$\$10{,}000{,}000 \times \frac{0.241}{100} = \$24{,}100 \tag{4.3}$$

To generalize the discussion here, let P denote the price of a fixed income instrument and let y denote the single factor that determines rate changes across the term structure, which, in the context of this chapter, is the 30-year HQM par rate. Finally, let Δy denote the change in the factor y and ΔP the change in the price given that change in y. With this notation, the DV01 of a fixed income instrument is estimated by,

$$DV01 \approx -\frac{1}{10{,}000}\frac{\Delta P}{\Delta y} \tag{4.4}$$

The slope of the NSC bond's price–rate curve at the current rate is estimated in Equations (4.1) and (4.2) with prices at levels one basis point below and one basis point above the current rate. A more accurate estimate might be obtained by shifting the current rate up and down by 0.5 basis points, or 0.1 basis points, etc. In the calculus, the limit of the resulting estimates of the slope, as the size of the shift shrinks to zero, is known as the *derivative* and denoted dP/dy. In some cases, like the yield-based metrics discussed later in this chapter, the derivative of the price–rate function can be written in closed form, that is, with a relatively simple mathematical formula. More generally, however, prices and the slope of the price–rate function have to be calculated numerically. In any case, the limit of Equation (4.4) as the size of the shift shrinks to zero gives DV01 in terms of the derivative instead of the slope,

$$DV01 = -\frac{1}{10{,}000}\frac{dP}{dy} \tag{4.5}$$

The derivative of a curve at a particular point is often illustrated with *tangent* lines. In Figure 4.3, the solid black line is the price–rate curve of the NSC century bonds, which is also shown in Figure 4.2. The dashed line is the tangent line of the curve at the current market rate. This means that the slope of the line equals the derivative of the curve at the current market rate, and that the line just touches the curve at that rate. Figure 4.3 also shows the tangent line to the curve at a rate 150 basis points below the current market rate. The slope of this tangent line is clearly steeper than the other, which means that the NSC 4.10s of 05/15/2121 are more sensitive to rates at 150 basis points below the current rate than they are at the current rate.

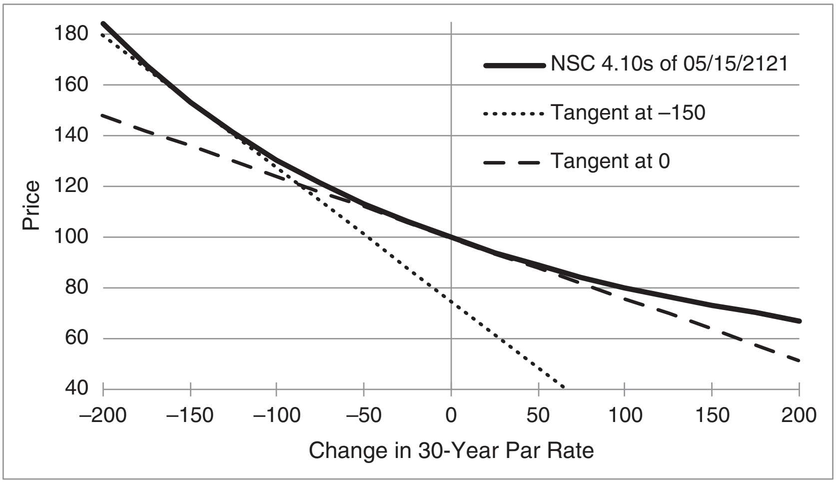

FIGURE 4.3 Tangent Lines to the Price–Rate Curve of the NSC 4.10s of 05/15/2121, as of Mid-May 2021.

DV01 is said to be a *local* measure of interest rate sensitivity because the slope of the price–rate curve, and, therefore, DV01, change as rate changes. This property of price–rate curves is known as convexity and is introduced later in the chapter.

This section concludes by noting that the DV01 of a portfolio of fixed income instruments is equal to the sum of the DV01s of its component instruments. A portfolio with 100 face amount of the NSC bonds – with a DV01 of 0.241 – and 500 face amount of the Treasury 1.625s of 05/15/2026 – with a DV01 of $5 \times 0.040 = 0.200$ – has a total DV01 of $0.241 + 0.200$, or 0.441. This very intuitive rule is proved formally in Appendix A4.1.

4.3 HEDGING A CENTURY BOND: PART I

Say that a market maker buys from a client \$10,000,000 face amount of the NSC 4.10s of 05/15/2121 at the going bid price. The market maker does not turn around and immediately sell those bonds at market, because that would likely mean selling at the same bid price, leaving no overall profit from the trades. Instead, the market maker waits until other clients appear, who are willing to pay the higher ask price. In this way, the market maker earns the bid–ask spread for providing immediacy or liquidity both to the client who originally sells the bonds and to the clients who later buy the bonds.

This strategy, however, exposes the market maker to the risk that the price of the corporate bond falls before it can be sold. The typical solution is to hedge that risk by selling a liquid Treasury bond when buying the corporate, and then buying back that Treasury bond when selling the corporate. Because a liquid Treasury bond, by definition, has a very narrow bid–ask spread, this strategy protects the market maker against falling prices at the cost of that narrow bid–ask spread, which leaves much of the wider, corporate bid–ask spread as profit.

The next question, therefore, is which Treasury bond to sell. Because, as mentioned earlier, rates of different terms can behave differently, a reasonable choice is to sell a Treasury bond with about the same maturity as the corporate bond being hedged. In the particular situation at hand, however, there is no Treasury bond with a maturity anywhere near 100 years. The best the market maker can do, therefore, is to sell one of the longest maturity Treasuries outstanding, like the 1.625s of 11/15/2050.

The final question, then, is what face amount of Treasury bonds to sell against the purchase of \$10 million face amount of the NSC bonds. One common solution is to ensure that the net DV01 of the combined position is zero. In other words, choose the face amount of Treasuries such that, if rates change by one basis point, the value of the net position is unchanged. Denoting the face amount of the Treasury hedge by F, and using the DV01s of the two bonds computed in Table 4.2, F is determined by the following equation,

$$F\frac{0.199}{100} + \$10{,}000{,}000\frac{0.241}{100} = 0 \tag{4.6}$$

The first term of (4.6) gives the change in the value of F face amount of Treasury bonds if rates fall by one basis point – 19.9 cents per 100 face amount, and, therefore, F times 0.199/100 for F face amount. The second term, along the same lines, gives the change in the value of the \$10 million face amount of NSC bonds if rates fall by one basis point. The equation as a whole, therefore, requires that the combined, hedged position neither gains nor loses value when rates fall by one basis point. And because the hedged profit and loss (P&L) when rates fall by some other number of basis points is just that number of basis points times the left-hand side of (4.6), the equation holds for that number of basis points as well (subject, of course, to the limitation of DV01 as a local measure of interest rate sensitivity).

Solving Equation (4.6) for F,

$$F = -\$10{,}000{,}000\frac{0.241}{0.199} = -\$12{,}110{,}553 \tag{4.7}$$

Hence, the market maker can hedge its \$10 million face amount of the NSC bonds by selling about \$12.1 million face amount of the Treasury bond. Intuitively, because the price of the NSC bonds changes by 24.1 cents per

100 face amount per basis point, while the Treasury bonds change by only 19.9 cents, the market maker has to sell a larger face amount of Treasuries than it buys of NSC bonds to achieve a DV01-neutral position.

To elaborate on how the hedge works, say that rates rise by five basis points after the market maker established the position. The NSC bonds fall in value, losing approximately $\$10{,}000{,}000 \times (0.241/100) \times 5 = \$120{,}500$. At the same time, the Treasury bonds also fall in value, gaining – because the market maker is short – approximately $\$12{,}110{,}553 \times (0.199/100) \times 5 = \$120{,}500$. Hence, as intended, the overall, hedged position does not gain or lose money.

In general, if the DV01s of bonds A and B are $DV01^A$ and $DV01^B$, then F^A face amount of bond A is hedged with F^B face amount of bond B such that,

$$F^A \frac{DV01^A}{100} + F^B \frac{DV01^B}{100} = 0 \tag{4.8}$$

$$F^B = -F^A \frac{DV01^A}{DV01^B} \tag{4.9}$$

Equation (4.8) is the generalization of Equation (4.6). The general solution, Equation (4.9), reveals two intuitive points about DV01 hedging. First, a long position in bond A is hedged by a short position in bond B. Second, the bond with the higher DV01 is traded in smaller quantity. In the market making example, a long position in the NSC bonds is hedged by a short position in Treasury bonds, and, because the DV01 of the NSC bonds is greater than that of the Treasury bonds, \$10 million of NSC bonds is hedged by \$12.1 million of Treasury bonds.

The section concludes with a reminder of the assumptions behind the DV01 hedge constructed here. First, rate shifts across terms are assumed to be proportional to those illustrated in Figure 4.1. If it turns out that the 100-year and 29.5-year par rates move differently than as assumed, the hedge will not work as intended. To the extent that this is a concern, the hedges described in Chapters 5 and 6 might be more appropriate. Second, the spreads of the NSC and Treasury bonds to the base curve are assumed to be constant. If it turns out that these spreads behave differently, then, again, the hedge will not work as intended. This concern is not as easily addressed. A safer hedge would be to sell a corporate bond, whose spread is more highly correlated with the NSC bond spread than is the Treasury bond spread. But hedging with a relatively illiquid corporate bond instead of an extremely liquid Treasury bond is likely to wipe out any market making P&L. Therefore, bearing spread risk for short periods of time may be an integral part of corporate bond market making and, in turn, may enter into the determination of bid–ask spreads in that market.

4.4 DURATION

Another popular metric for interest rate sensitivity is *duration*. Whereas DV01 measures the change in price for a change in rates, duration measures the percentage change in price for a change in rates. Like DV01, duration can be defined in any one-factor framework, but practitioners often use the term duration to mean yield-based duration (described in Section 4.7) and use the term *effective duration* to mean the more general case presented here.

Using the same notation as in Section 4.2, duration, D, is estimated in terms of the slope as,

$$D \approx -\frac{\Delta P/P}{\Delta y} = -\frac{1}{P}\frac{\Delta P}{\Delta y} \tag{4.10}$$

and given in terms of the derivative as,

$$D = -\frac{dP/P}{dy} = -\frac{1}{P}\frac{dP}{dy} \tag{4.11}$$

Because $\Delta P/P$ and dP/P both represent percentage change in price, Equations (4.10) and (4.11) express duration as the percentage change in price for a change in rates. Also, for intuition, it is useful to rewrite Equation (4.10) as,

$$\frac{\Delta P}{P} \approx -D\Delta y \tag{4.12}$$

Table 4.3 shows the calculation of duration for the three bonds introduced in Table 4.1. Estimating the duration of the NSC 4.10s of 05/15/2121 with Equation (4.10),

$$-\frac{1}{99.9390}\frac{99.6990 - 100.1801}{0.02\%} = 24.1 \tag{4.13}$$

Applying Equation (4.12), the percentage change in the price of the NSC bond for a decline in rates of 100 basis points, or 1%, is $-24.1 \times (-1\%) = 24.1\%$. Hence, a duration of 24.1 roughly means that a fall in rates of 100 basis points increases bond price by 24.1%. This interpretation is only roughly correct because duration, like DV01, is based on the slope of the price–rate curve and, therefore, is a local measure of price change.

TABLE 4.3 Calculating Duration for Bonds in Table 4.1, as of Mid-May 2021.

Bond	Price −1bp	Price	Price +1bp	Slope	Duration
NSC of May 2121	100.1801	99.9390	99.6990	−2,405.5	24.1
Treasury of May 2026	103.9621	103.9219	103.8817	−402.00	3.9
Treasury of Nov. 2050	84.5899	84.3906	84.1919	−1,990	23.6

While traders tend to rely on DV01, asset managers tend to rely on duration. As in the market-making example of the previous section, traders typically want to ensure that the dollar changes in the value of long and short positions offset each other. Also, because the sizes of their positions can fluctuate rapidly, and because they typically borrow money to buy bonds, they tend to focus on dollar P&L rather than on returns on fixed amounts of invested cash. Asset managers, in contrast, typically invest a slowly changing pool of funds and focus on rates of return. Looking at the durations in Table 4.3, an asset manager immediately sees that the NSC bond has the most interest rate risk, in that an investment in that bond loses 2.41% for a 10-basis point increase in rates. An investment in the Treasury 1.625s of 11/15/2050 has slightly less risk, losing 2.36% for that 10-basis-point increase in rates, while an investment in the Treasury 1.625s of 05/15/2026 has the least risk, losing only 0.39% in that scenario. The asset manager can weigh these risks against expected returns and other factors in making a final investment decision.

This section concludes by noting the duration of a portfolio is equal to the weighted sum of the durations of its component holdings, where the weights are percentages of portfolio value. For example, the duration of a portfolio with 25% of its value in the NSC bonds – with a duration of 24.1 – and 75% of its value in the Treasury 1.625s of 05/15/2026 – with a duration of 3.9 – is $25\% \times 24.1 + 75\% \times 3.9 = 8.95$. The DV01 of a portfolio is the sum of its component DV01s, while the duration of a portfolio is the value-weighted sum of its component durations, because DV01 represents a change in price, while duration represents a percentage change in price. A formal proof is in Appendix A4.1.

4.5 CONVEXITY

As can be seen from the tangent lines in Figure 4.3, the interest rate sensitivity of a bond falls as rates increase. Making this point more directly, Figure 4.4 graphs the DV01 of the Treasury 1.625s of 11/15/2050, a 29.5-year bond, and of the Treasury 1.625s of 05/15/2026, a five-year bond. DV01 falls as rates increase, but the rate of decline is much faster for the 29.5-year bond than for the five-year bond.

A curve is said to have a convex shape if a line connecting two points on that curve lies above the curve. The price–rate curves of coupon bonds are convex, and Figure 4.3 makes it clear that the convex shape of the price–rate curve is tantamount to DV01 falling as rates increase. For this reason, the relationship between DV01 and the level of rates is called *convexity*. The property of DV01 falling as rates increase is called *positive convexity*, while the property of DV01 rising as rates increase is called *negative convexity*.

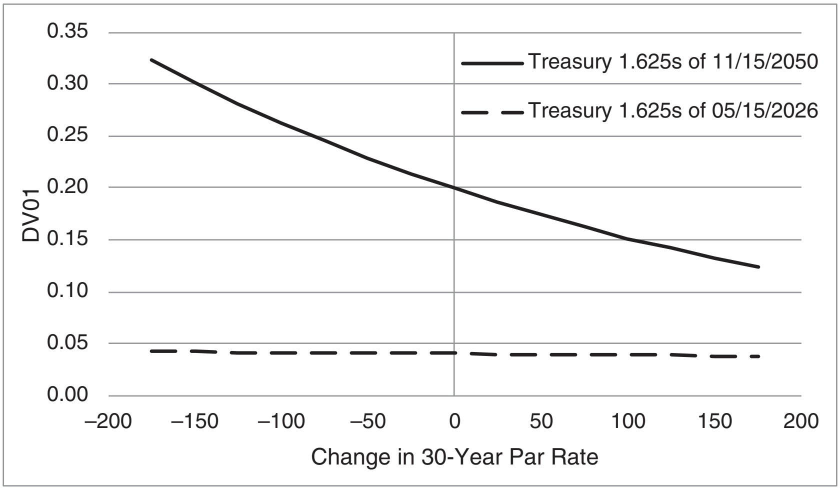

FIGURE 4.4 DV01s of the Treasury 1.625s of 11/15/2050 and of the Treasury 1.625s of 05/15/2026, as of Mid-May 2021.

Mathematically, convexity, C, is defined as,

$$C = \frac{1}{P}\frac{d^2P}{dy^2} \tag{4.14}$$

the second derivative of the price–rate function divided by price. To summarize, all of these statements are equivalent: the price–rate curve is convex; its second derivative is positive; its first derivative becomes less negative as rates increase; and its DV01 falls as rates increase.

Table 4.4 calculates the convexity of the NSC bond and of the two Treasury bonds as of mid-May 2021. For each bond, three prices are given: the current market price, the price after the 30-year par rate falls by one basis point, and the price after the 30-year par rate rises by one basis point. Then, two first derivatives, or slopes, are computed for each bond. These slopes are computed just as in Equation (4.1), but the slopes here are centered around the 30-year par rate plus and minus 0.5 basis points.

The final step in Table 4.4 is to estimate the convexity as defined in Equation (4.14). The second derivative centered at the current rate (i.e., a "change" of 0.0) is estimated as the change in the first derivatives from +0.5 to −0.5 basis points divided by the change in rates of +0.5 – (−0.5), or 1 basis point. Dividing the result by price gives the estimate of convexity centered at current market levels. For the NSC bond,

$$C = \frac{1}{99.939}\frac{-2{,}399.67 - (-2{,}410.67)}{0.01\%} = 1{,}101 \tag{4.15}$$

TABLE 4.4 Calculating Convexity for Bonds in Table 4.1, as of Mid-May 2021. Rate Changes Are in Basis Points.

Rate Change	Price	1st Derivative	Convexity
	NSC 4.10s of 05/15/2121		
−1.0	100.180067		
−0.5		−2,410.67	
0.0	99.939000		1,101
0.5		−2,399.67	
1.0	99.699033		
	Treasury 1.625s of 11/15/2050		
−1.0	84.589932		
−0.5		−1,993.08	
0.0	84.390624		648
0.5		−1,987.61	
1.0	84.191863		
	Treasury 1.625s of 05/15/2026		
−1.0	103.962050		
−0.5		−401.75	
0.0	103.921875		12
0.5		−401.63	
1.0	103.881712		

As expected from the curvatures of the price–rate curves in Figure 4.2, the convexity measure is largest for the NSC century bond, smaller for the 29.5-year Treasury bond, and yet smaller for the five-year Treasury bond.

Convexity values are not as easily interpreted as DV01 and duration values, but consider the following. Equation (4.12) showed that the percentage change in a bond's price approximately equals the negative of its duration times the change in rate. However, Appendix A4.2 shows that a better approximation uses both the bond's duration and convexity,

$$\frac{\Delta P}{P} \approx -D\Delta y + \frac{1}{2}C\Delta y^2 \tag{4.16}$$

Because duration appears in the first term of (4.16) and convexity in the second, using duration alone is called a *first-order approximation*, while using both duration and convexity is called a *second-order approximation*.

Illustrating with the NSC bonds, which have a duration of 24.1 (Table 4.3) and a convexity of 1,101 (Table 4.4), estimates of the percentage change in price after rates fall by 100 basis points, or 1%, are,

$$-24.1 \times (-1\%) = 24.1\% \tag{4.17}$$

using duration alone, and,

$$-24.1 \times (-1\%) + 0.5 \times 1{,}101 \times (-1\%)^2 = 29.6\% \tag{4.18}$$

using both duration and convexity. The actual price of the bond after rates decline by 100 basis points is 130.898, which translates into a true percentage price change of (130.898 – 99.939)/99.939, or 30.978%. Therefore, adding the convexity term in Equations (4.16) and (4.18) does result in a more accurate approximation.

Figure 4.5 makes the same point graphically. The actual price of the NSC bonds is given by the solid curve. The approximation to prices after various changes in rates using just duration or DV01 is given by the dashed line, which is the same dashed tangent shown in Figure 4.3. And the approximation using both duration and convexity is given by the dotted line. Because both duration and convexity are local estimates, using derivatives at current market levels to estimate prices, both the dashed and dotted lines are less and less accurate as rates move further from current levels. It is clear to the eye, however, that the approximation using both duration and convexity is relatively close to the true price for larger changes in rates than is the approximation using duration alone.

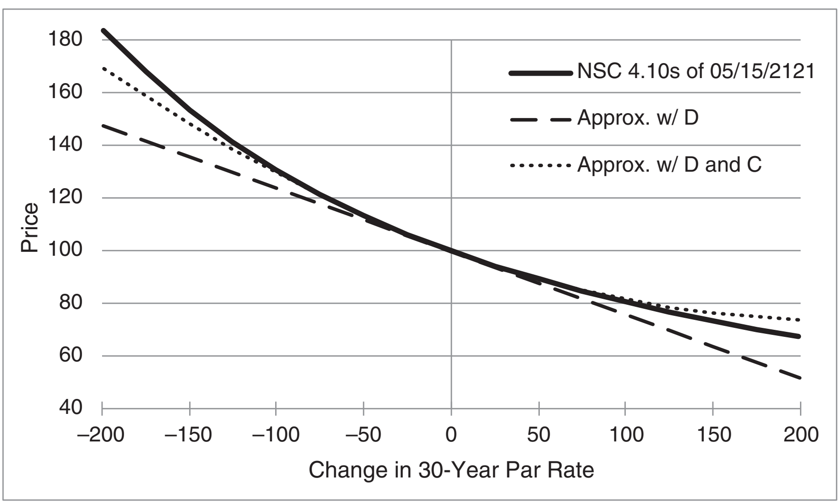

FIGURE 4.5 Price–Rate Curve of the NSC 4.10s of 05/15/2121, as of Mid-May 2021, with Price Approximations Using Duration and Using both Duration and Convexity.

Returning now to interpreting convexity numbers, the difference between approximating percentage price change with both duration and convexity [Equation (4.18)] and with duration alone [Equation (4.17)] is the term,

$$0.5 \times 1{,}101 \times (1\%)^2 = 5.5\% \tag{4.19}$$

Therefore, noting that $(1\%)^2$ equals one basis point, the correction of 5.5% in (4.19) is half the convexity divided by 10,000. In other words, for a 1% change in rates, the second-order correction to the duration approximation of price change is half the convexity divided by 10,000 (i.e., 0.055), or equivalently, the percentage correction is half the convexity divided by 100 (i.e., 5.5).

The section concludes by noting that the convexity of a portfolio equals the weighted sum of the convexities of its component holdings, where the weights are percentages of portfolio value. The convexity of a portfolio with 25% of its value in the NSC bonds – with a convexity of 1,101 – and 75% of its value in the Treasury 1.625s of 05/15/2026 – with a convexity of 12 – is $25\% \times 1{,}101 + 75\% \times 12 = 284.25$. Portfolio convexity is a weighted sum, as is portfolio duration, because both metrics divide by price: duration divides the change in price by price, and convexity divides the second derivative divided by price. A formal proof is given in Appendix A4.1.

4.6 HEDGING A CENTURY BOND: PART II

Section 4.3 explained why and how a market maker hedges the DV01 from \$10 million face amount of the NSC 4.10s of 05/15/2121 by selling \$12.1 million face amount of the Treasury 1.625s of 11/15/2050. This section shows that this hedge leaves the market maker with a *long convexity* position and describes the resulting risk and P&L implications.

Figure 4.6 shows the P&L of \$100 million face amount of the NSC century bond and of \$121 million face amount of the 29.5-year Treasury bond. For example, the prices of these bonds in mid-May 2021 were 99.939 and 84.391, respectively, and if rates fell by 100 basis points, their prices would be 130.898 and 107.309, respectively. Hence, the P&L from \$100 million of the one and \$121 million of the other is \$100 million × (130.898 – 99.939)/100, or \$30.96 million, and \$121 million × (107.309 – 84.391)/100, or \$27.7 million, respectively.

Section 4.3 showed that, with these face amounts, the two positions have the same DV01 at current market levels. Figure 4.6 reflects this fact in that the two P&L curves are tangent to each other, that is, they have the same slope, at current rates. Both bonds are positively convex, of course, but, because the NSC bonds are more convex (Table 4.4), their P&L exceeds the P&L on the Treasury bonds whether rates fall or rise. This result is emphasized by the dotted line in the figure, graphed against the right axis, which is

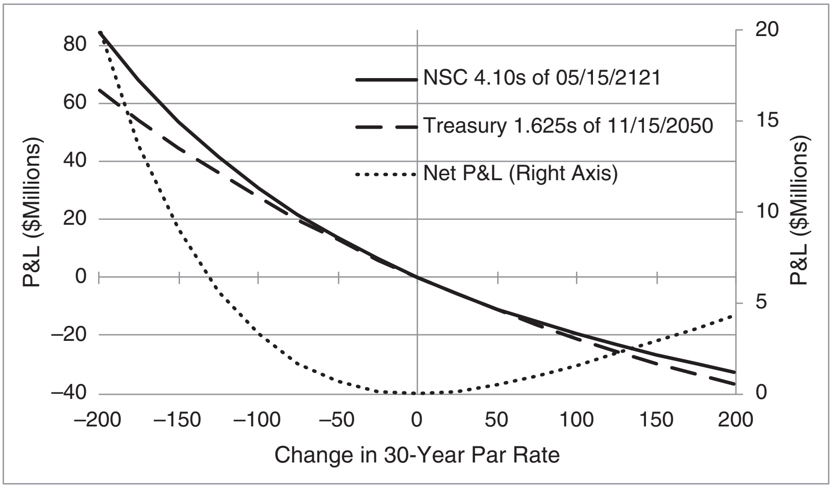

FIGURE 4.6 P&L of a Long Position of $100 Million Face Amount of the NSC 4.10s of 05/15/2121, a DV01-Equivalent Long Position in the Treasury 1.625s of 11/15/2050, and a Position Long the NSC Bonds and Short the Treasury Bonds, as of Mid-May 2021.

the P&L of the NSC bonds minus the P&L of the Treasury bonds and which, in turn, is the net P&L of the market maker in this application. This net P&L is zero, of course, at current market levels but is positive for both negative and positive rate changes, and particularly large for large, negative changes.

Expanding on the results of the previous paragraph, for unchanged rates and prices, the P&L of both bond holdings is clearly zero. If rates fall, both bonds increase in price and – because they are both positively convex – both their DV01s increase as well. But because the NSC bonds are more convex, their DV01 increases by more, which, for falling rates, means higher profits. Instead, if rates rise, both bonds decrease in price and in DV01. But because the NSC bonds are more convex, their DV01 falls by more, which, for rising rates, means lower losses. Therefore, whether rates fall or rise, the P&L on the NSC bonds exceeds the P&L on a DV01-equivalent amount of Treasury bonds.

The market maker is said to have a positively convex position and to be long convexity because the convexity of its long position (NSC bonds) is greater than the convexity of its short position (Treasury bonds) and because its P&L seems to be positive whether rates fall or rise. But is it really the case that the market maker profits from the position no matter what? If so, why don't arbitrageurs initiate positions that are long the century NSC bonds and short the 29.5-year Treasury bonds?

These questions can be answered by noting that Figure 4.6 omits an important variable: time. As mentioned earlier, all of the figures in this chapter assume an instantaneous change in rates, that is, with no passage of time. As it turns out, however, bonds or portfolios that are more convex tend to earn less over time. Therefore, the full story of a positively convex position, or of being long convexity, is that the position profits if rates move down or up by a sufficient amount. If rates change by less than that, or stay the same, the position loses money. In this sense, a long convexity position is also long volatility. In any case, this is a fundamental property of asset pricing, applicable to both stock options and coupon bonds. Section 4.8 explores the topic further in the context of pension asset–liability management, and Chapter 8 shows how convexity fits into a general framework of returns.

4.7 YIELD-BASED DV01, DURATION, AND CONVEXITY

Previous sections described DV01, duration, and convexity in a general, one-factor framework: specify the way changes in a rate or factor change the entire term structure, re-price bonds after the change, and compute risk metrics. This section is about yield-based metrics, for which a change in rates means a fixed change in bond yields. These metrics have two significant weaknesses. One, they are defined only for bonds with fixed cash flows, but not, for example, for a callable bond or a mortgage. Two, their use implicitly assumes parallel shifts in bond yields, which is not an empirically sound assumption. Nevertheless, there are several reasons to study and understand yield-based metrics. First, they are simple to compute and easy to understand, and, in many situations, are perfectly reasonable to use. Second, these metrics are widely used across the financial industry. Third, much of the intuition gained from understanding these metrics carries over to more general frameworks.

The reason that yield-based metrics are easy to compute is that price can be written as a function of yield, in the forms of Equations (3.7) and (3.8). Recalling that c is the annual coupon payment and T is the number of years to maturity, these equations are repeated here for convenience,

$$P = \frac{c}{2}\sum_{t=1}^{2T}\frac{1}{\left(1+\frac{y}{2}\right)^t} + \frac{100}{\left(1+\frac{y}{2}\right)^{2T}} \tag{4.20}$$

$$P = \frac{c}{y}\left(1 - \frac{1}{\left(1+\frac{y}{2}\right)^{2T}}\right) + \frac{100}{\left(1+\frac{y}{2}\right)^{2T}} \tag{4.21}$$

With these expressions, DV01 and duration, as defined in Equations (4.5) and (4.11), can be solved explicitly. Calculate the derivative of the previous price-yield relationships and then divide by $-10{,}000$ to find DV01 or by $-P$ to find duration. The resulting formulas for yield-based DV01 are,

$$DV01 = \frac{1}{10{,}000}\frac{1}{1+\frac{y}{2}}\left[\frac{c}{2}\sum_{t=1}^{2T}\frac{t}{2}\frac{1}{\left(1+\frac{y}{2}\right)^{t}} + T\frac{100}{\left(1+\frac{y}{2}\right)^{2T}}\right] \tag{4.22}$$

$$DV01 = \frac{1}{10{,}000}\left[\frac{c}{y^2}\left(1-\frac{1}{\left(1+\frac{y}{2}\right)^{2T}}\right) + T\left(1-\frac{c}{100y}\right)\frac{100}{\left(1+\frac{y}{2}\right)^{2T+1}}\right] \tag{4.23}$$

and for yield-based duration are,

$$D = \frac{1}{P}\frac{1}{1+\frac{y}{2}}\left[\frac{c}{2}\sum_{t=1}^{2T}\frac{t}{2}\frac{1}{\left(1+\frac{y}{2}\right)^{t}} + T\frac{100}{\left(1+\frac{y}{2}\right)^{2T}}\right] \tag{4.24}$$

$$D = \frac{1}{P}\left[\frac{c}{y^2}\left(1-\frac{1}{\left(1+\frac{y}{2}\right)^{2T}}\right) + T\left(1-\frac{c}{100y}\right)\frac{100}{\left(1+\frac{y}{2}\right)^{2T+1}}\right] \tag{4.25}$$

This formulation of yield-based duration is also known in the industry as *modified* and *adjusted* duration.[1]

The terms inside the square brackets of Equations (4.22) and (4.24) have an intuitive interpretation that sheds light on these risk metrics. The terms $t/2$ and T represent the times to receipt of each cash flow, that is, 0.5 years, 1 year, 1.5 years, and so forth, out to T years, the years to maturity. And each of these terms is multiplied by the present value of the cash flow to be received at that time, that is, $(c/2)/(1+y/2)^t$ for the coupon payments and $100/(1+y/2)^{2T}$ for the principal payment. Putting this all together, the sum of the terms in brackets can be described as the weighted sum of the times at which cash flows are received, with each weight equal to the present value of the cash flow received at that time. Furthermore, in the case of duration,

[1]This terminology is a historical artifact. The first metric of this sort was *Macaulay duration*, which equals the expressions in (4.24) or (4.25) $\times\ 1+y/2$. Over time, however, the definitions in the text became the industry standard. Macaulay duration has the advantage that the resulting duration of a zero coupon bond is exactly equal to its years to maturity. The definitions in the text, however, have the advantage that they are exactly equal to the percentage change in price for a change in rates.

where the bond price can be moved inside the brackets, each weight can be described as the present value of the cash flow received at that time divided by the bond price, which is the sum of all the present values. Therefore, this weight is simply the proportion of the bond's value contributed by that cash flow.

Table 4.5 illustrates this interpretation by calculating the yield-based DV01 and duration of the Treasury 1.625s of 05/15/2026 as of mid-May 2021. The first column gives the term of each cash flow; the second column gives the cash flows; and the third column gives the present value of each cash flow discounted at 0.82277%, the bond's yield at the time. For example, the present value of the coupon payable in 2.5 years is,

$$\frac{0.8125}{\left(1+\frac{0.82277\%}{2}\right)^{5}} = 0.7960 \tag{4.26}$$

The sum of the present values of the cash flows is just the market price, in this case, 103.9219. The fourth column gives the present value of each cash flow as a percentage of the market price. Not surprisingly, for a bond maturing in five years and trading at a small premium, the principal accounts for a large share of the bond's value, in this case, over 93%.

TABLE 4.5 Calculating the Yield-Based DV01 and Duration of the Treasury 1.625s of 05/15/2026 at a Yield of 0.82277%, as of Mid-May 2021.

(1) Term	(2) Cash Flow	(3) Present Value	(4) PV/Price (%)	(5) Term × PV/Price	(6) Term-Wtd PV
0.5	0.8125	0.8092	0.779	0.0039	0.4046
1.0	0.8125	0.8059	0.775	0.0078	0.8059
1.5	0.8125	0.8026	0.772	0.0116	1.2038
2.0	0.8125	0.7993	0.769	0.0154	1.5985
2.5	0.8125	0.7960	0.766	0.0191	1.9900
3.0	0.8125	0.7927	0.763	0.0229	2.3782
3.5	0.8125	0.7895	0.760	0.0266	2.7632
4.0	0.8125	0.7862	0.757	0.0303	3.1450
4.5	0.8125	0.7830	0.753	0.0339	3.5236
5.0	100.8125	96.7575	93.106	4.6553	483.7877
Sum		103.9219	100.000	4.8267	501.6005
Duration:				4.8069	
DV01:					0.0500

The fifth column of the table multiplies each present value proportion by its term: 0.5 weighted by 0.779%, 1.0 weighted by 0.775%, and so forth, and 5.0 weighted by 93.106%. In this way, the sum of the column is the weighted average of the payment times, with weights equal to the value proportion of each payment. In terms of Equation (4.24), this weighted average, 4.8267, is the sum of the terms in brackets divided by price. Duration, therefore, is 4.8267/(1 + 0.82277%/2), or 4.8069. This interpretation explains why many market participants say that this bond has a duration of 4.83 *years*: the value of the bond is paid, on average, in 4.83 years.[2] This interpretation also explains why the duration of a coupon bond is somewhat less than its maturity: while most of a bond's value is paid at maturity, some portion is paid earlier.

Lastly, the sixth column of the table gives term times present value. The sum of this column is the sum of the terms in brackets in Equation (4.22), which means that DV01 is $(1/10{,}000) \times 501.6005/(1 + 0.82277\%/2) = 0.0500$.

As mentioned earlier, the intuition derived from yield-based measures is often useful for understanding more general risk metrics. A lot of this intuition arises from the relatively simple expressions for yield-based DV01 and duration in Equations (4.23) and (4.25), and from the extremely simple expressions in the case of zero coupon bonds and par bonds. To derive these latter expressions, simply substitute $c = 0$ and then $c = 100y$ into the price equation, (4.21), and into the DV01 and duration equations just referenced, to obtain the following. For a zero coupon bond,

$$DV01_{c=0} = \frac{T}{100\left(1+\frac{y}{2}\right)^{2T+1}} \tag{4.27}$$

$$D_{c=0} = \frac{T}{\left(1+\frac{y}{2}\right)} \tag{4.28}$$

And for a par bond,

$$DV01_{c=100y} = \frac{1}{100y}\left(1-\frac{1}{\left(1+\frac{y}{2}\right)^{2T}}\right) \tag{4.29}$$

$$D_{c=100y} = \frac{1}{y}\left(1-\frac{1}{\left(1+\frac{y}{2}\right)^{2T}}\right) \tag{4.30}$$

[2]A purely mathematical reason that the units of duration is years is that duration is a percentage change divided by a change in interest rate *per year*.

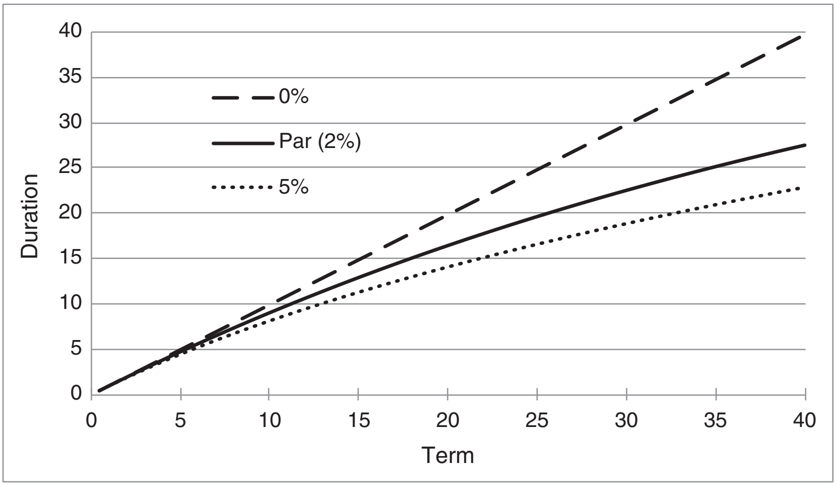

FIGURE 4.7 Yield-Based Duration for Bonds with Coupons of 0%, 2%, and 5%. Yield Equals 2%.

The discussion now turns to using these simple expressions to understand how yield-based duration and DV01 depend on a bond's maturity, coupon, and yield. Figure 4.7, fixing yield for all bonds at 2%, graphs the duration of bonds of terms out to 40 years with coupons of 0%, 2%, and 5%. Several lessons can be taken from this figure.

First, the duration of a zero coupon bond is approximately equal to its term, which can also be seen from Equation (4.28). Second, the duration of a par bond, in this case, of all bonds with a coupon of 2%, increases with term, which can also be deduced from Equation (4.30).[3] In addition, however, the figure illustrates how the duration of par bonds increases less than linearly with term. Out to terms of five years, duration approximately equals term, but, for longer terms, duration falls more and more below term. The duration of a 10-year par bond is about 9.0; of a 20-year bond, 16.4; of a 30-year bond, 22.5; and of a 40-year bond, 27.4. Third, the duration of a premium bond, with a coupon of 5%, is less than the duration of a par bond. More generally, looking at the figure as a whole, bonds with higher coupons have lower durations. The intuition here is that higher-coupon bonds pay a

[3]While usually the case, it is not true that duration of any bond always increases with term. The durations of deeply discounted coupon bonds increase with term, arbitrarily close to the zero coupon line in Figure 4.7, but can then decline, with further increases in term, as the duration falls to that of a perpetuity. The duration of a perpetuity is the reciprocal of the yield, which can be shown letting T approach infinity in Equations (4.21) and (4.25).

greater fraction of their value earlier, which, in turn, means that the lower terms are more heavily weighted in the calculation of duration. Put another way, higher-coupon bonds are effectively shorter-term bonds and, therefore, have lower durations.

Figure (4.8) graphs the DV01s of the same set of bonds, continuing to hold the yields of all bonds equal to 2%. The DV01s of par bonds increase with maturity, which follows directly from inspection of Equation (4.29). As with duration, however, the increase is less than linear. The DV01s of par bonds with terms of five years or less are about equal to term divided by 100. The DV01 of a five-year bond at a yield of 2%, for example, is 0.047, which is slightly less than 0.05, or 5 cents per 100 face amount. But, for longer terms, the difference is larger: the DV01 of the 10-year par bond in the figure is 9 cents; of the 20-year bond, 16 cents; of the 30-year bond, 22 cents; and of the 40-year bond, 27 cents.

Unlike duration, however, Figure (4.8) shows that DV01 increases with coupon. To understand this, combine the definitions of DV01 and duration in Equations (4.5) and (4.11) to see that,

$$DV01 = \frac{P \times D}{10{,}000} \tag{4.31}$$

In thinking about how DV01 changes with term, therefore, there is a price effect in addition to a duration effect. The duration effect almost always causes DV01 to increase with term, along the lines of the previous discussion. The price effect, however, can reinforce or counter this duration effect. For par bonds, whose prices are always 100, there is no price effect. For premium

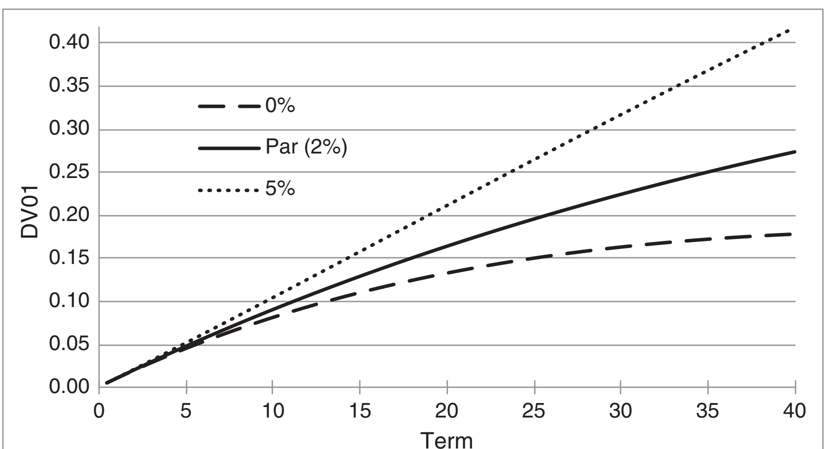

FIGURE 4.8 Yield-Based DV01 for Bonds with Coupons of 0%, 2%, and 5%. Yield Equals 2%.

bonds, whose prices increase with term (recall Figure 3.1), the price effect reinforces the duration effect and, therefore, as shown in Figure (4.8), the DV01s of 5% bonds increase much more rapidly with term than do the DV01s of par bonds. For discount bonds, by contrast, whose prices decrease with term, the price effect works against the duration effect and, therefore, as shown in the figure, the DV01s of zero coupon bonds increase more slowly with term than those of par bonds. In fact, inspection of Equation (4.27) reveals that, at large enough terms, the DV01s of zero coupon bonds fall as term increases further, that is, the price effect, $(1+y/2)^{2T+1}$, eventually dominates the duration effect, T.

Having described how DV01 and duration vary with term and coupon rate, the discussion turns to the effect of yield. It is clear from Equation (4.22) that DV01 falls as yield increases. This fact was introduced earlier, as an implication of the convex shape of the price–rate curve. As it turns out, increasing yield also lowers duration. Intuitively, increasing yield lowers the present value of all payments, but lowers the present value of the longer payments the most. This, in turn, lowers the proportions of bond value in the longer payments, lowers their weights in the duration calculation, and, therefore, lowers the duration of the bond.

Figure 4.9 illustrates how duration changes with yield, graphing the duration of par bonds of various terms at yields of 0.5%, 2%, and 5%. As just discussed, duration is lower at higher yields, significantly so for longer terms. Furthermore, the difference between duration and term, discussed earlier, is greater at higher yields. At a yield of 0.5%, the durations of 10- and 30-year par bonds are 9.7 and 27.8, respectively, while, at a yield of 5%, those durations are 7.8 and 15.5, respectively.

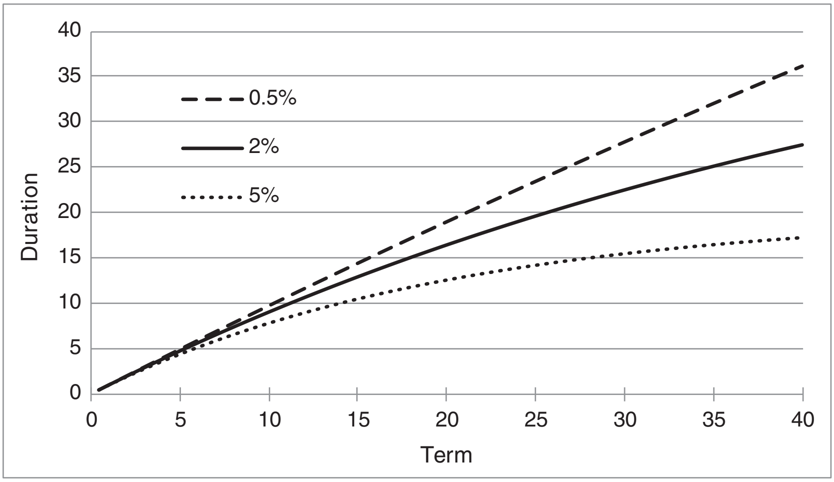

FIGURE 4.9 Duration of Par Bonds, with Yields Equal to 0.5%, 2%, and 5%.

The section turns now to yield-based convexity. Given the general definition of convexity in Equation (4.14), an expression for yield-based convexity can be found by taking the second derivative of Equation (4.20) and dividing by price. The resulting formula is,

$$C = \frac{1}{P\left(1+\frac{y}{2}\right)^2}\left[\frac{c}{2}\sum_{t=1}^{2T}\frac{(t/2)((t+0.5)/2)}{\left(1+\frac{y}{2}\right)^t} + \frac{100T(T+0.5)}{\left(1+\frac{y}{2}\right)^{2T}}\right] \tag{4.32}$$

As in the formula for yield-based duration, Equation (4.24), the terms inside the brackets of (4.32) multiply the present value of payments by some function of the term of those payments. The big difference between the two, however, is that, in the duration formula, that function of term is linear (i.e., t and T), while in the convexity formula, the function is quadratic, $(t/2)((t+0.5)/2)$ and $T(T+0.5)$. The implication is that convexity increases much faster with term than duration. This can be seen in the more general case from Tables 4.3 and 4.4. The durations of the 5- and 29.5-year Treasuries, and the 100-year NSC bonds are 3.9, 23.6, and 24.1, respectively, while their convexities are 12, 648, and 1,101, respectively.

For completeness, the convexity formulas for zero coupon and par bonds are included here. For zero coupon bonds,

$$C_{c=0} = \frac{T(T+0.5)}{\left(1+\frac{y}{2}\right)^2} \tag{4.33}$$

and for par bonds,

$$C_{c=y} = \frac{2}{y^2}\left[1 - \frac{1}{\left(1+\frac{y}{2}\right)^{2T}}\right] - \frac{2T}{y\left(1+\frac{y}{2}\right)^{2T+1}} \tag{4.34}$$

The section concludes with a note about the assumptions underlying the use of yield-based metrics. The yield-based DV01, duration, and convexity for each bond are determined by shifting the yield of that bond. Therefore, any risk management strategy across bonds implicitly assumes that yields of all included bonds move up and down together, in parallel. For example, if one bond has a yield-based DV01 of 0.05 and a second of 0.10, then hedging 100 face amount of the second by selling 200 face amount of the first is perfectly successful only if the yields of both bonds move by the same amount. If a trader hedges a 9.5-year bond with a 10-year bond over a short period of time, the assumption of parallel shifts in yield might be reasonable enough. Hedging a two-year bond with a 10-year bond, however, assuming that those yields move in parallel, is not likely to produce the desired, hedged outcome.

4.8 THE BARBELL VERSUS THE BULLET

This section presents a stylized example of asset–liability management for defined-benefit pension liabilities; applies the concepts of duration and convexity to hedging in this context, and explains how a choice arises between *barbell* and *bullet* asset portfolios, or, more generally, across asset portfolios with different convexities.

Figure 4.10 shows the liabilities of a stylized defined-benefit pension plan. Each black dot represents the total expected payment of the fund to retirees over a six-month period. The first total payment, in six months, is \$2 million. Subsequent total payments increase, as more existing employees retire and become eligible for benefits, reaching a peak of \$4 million in 10 years. From then on, deaths of retirees gradually reduce total payments, which fall to zero in 60 years. The gray dots represent the present value of each total payment, where liabilities are discounted at the HQM curve as of mid-May 2021.

The total present value of the liabilities in the figure is about \$140 million. For simplicity, assume that this pension fund is fully funded, so that the pension fund has \$140 million cash on hand. The asset–liability management problem of the fund, therefore, is to invest the \$140 million so as to be able

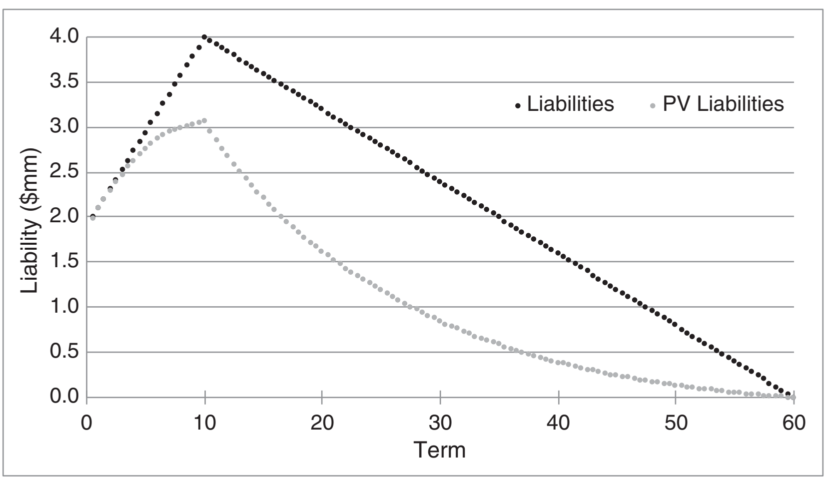

FIGURE 4.10 Liabilities of a Stylized Defined-Benefit Pension Fund. Present Values Are Computed by Discounting at the HQM Curve, as of May 2021.

to make good on pension liabilities over time, or even to earn some excess over those liabilities. Since discounting is done at the HQM corporate curve, the pension fund would break even by realizing investment returns equal to HQM rates. If the pension fund earned lower rates, likely, for example, if it were to invest in Treasuries, then there would not be enough cash available over the 60-year horizon to meet all of the pension obligations. And if the pension earned higher rates, perhaps by investing in stocks, then it would have more than enough to fulfill its pension obligations. In this case, however, there would be a significant risk that stocks perform relatively poorly over the period and leave the pension without enough funds to meet its obligations.

Because many pension funds make future payment commitments in line with returns on corporate bonds, the considerations in the previous paragraph lead many funds to invest a significant portion of their funds in corporate bonds. For the purposes of this section, it is assumed that all of the $140 million of pension assets are invested in corporate bonds. More specifically, it is assumed that the pension fund can invest in one or two of the Johnson & Johnson (JNJ) bonds listed in Table 4.6. As mentioned in Chapter 3, all of these bonds were issued by JNJ in August 2020. For expository purposes, these bonds are referred to by their original terms, which are not that different from their terms as of mid-May 2021. For example, the 0.55s of 09/1/2025 are called the five-year bonds, while the 2.45s of 09/1/2060 are called the 40-year bonds. The durations and convexities in the table are all yield-based metrics.

The duration and convexity of the pension liabilities, computed by a parallel shift in the HQM par rate curve, are 16.66 and 211.37, respectively. To hedge against the risk that rates fall, which would increase the present value of the liabilities, the pension fund decides to invest in an asset portfolio with a duration of 16.66, the duration of the liabilities. But in which of the JNJ bonds should the fund invest?

Perhaps the most straightforward answer is to invest in the 20-year bonds, which have a duration of 15.56. With its assets invested in this bond,

TABLE 4.6 Selected Johnson & Johnson Bond Issues. Yields, Durations, and Convexities Are as of Mid-May 2021. Coupons and Yields Are in Percent.

Coupon	Maturity	Yield	Duration	Convexity
0.55	09/1/2025	0.717	4.23	10.06
0.95	09/1/2027	1.238	6.08	20.33
1.30	09/1/2030	1.846	8.69	41.31
2.10	09/1/2040	2.686	15.56	141.49
2.25	09/1/2050	2.849	20.72	269.41
2.45	09/1/2060	2.962	24.09	393.17

the fund would have an asset duration equal to that of the bond, 15.56, and a liability duration, given already, of 16.66. Were rates to fall by 100 basis points, then, along the lines of previous sections, assets would increase in value by approximately 15.56% and liabilities by approximately 16.66%. The result is close to hedged, but not quite. On the initial portfolio value of \$140 million, the present value of the liabilities would then exceed the value of the assets by (16.66% – 15.56%) × \$140 million, or 1.1% of value, or \$1,540,000.

To clean up that hedge, the pension fund might use cash. Denote the amount held in cash by M, so that the fund's remaining funds, \$140 – M, are invested in the 20-year bond. Then, as mentioned in Section 4.4, the duration of this portfolio is the weighted average of its component durations. Note, too, that the duration of cash equals zero: its present value always equals its amount, no matter how rates change. Therefore, setting the duration of this asset portfolio equal to the duration of the liabilities,

$$\frac{M}{140} \times 0 + \frac{140 - M}{140} \times 15.56 = 16.66 \tag{4.35}$$

Solving, $M = -9.90$, meaning that the pension fund would borrow \$9.90 million, or about 7% of the \$140 million, and invest \$140 – (–9.90) = \$149.90 or about 107% in the 20-year bond.

The asset portfolio just described hedges the duration risk of the liabilities. But how does the convexity of the asset portfolio compare with the convexity of the liabilities? As mentioned in Section 4.5, the convexity of a portfolio is the weighted average of the convexity of its components. Here, with the convexity of cash equal to zero and the convexity of the 20-year bond equal to 141.49, the convexity of the asset portfolio is,

$$\frac{-9.90}{140} \times 0 + \frac{149.90}{140} \times 141.49 = 151.50 \tag{4.36}$$

which is less than the convexity of the liabilities, given earlier as 211.37. Hence, this asset portfolio results in the pension fund having an overall negatively convex position. Figure 4.11 graphs the P&L of this asset portfolio and of the liabilities, where the P&L of the bond is calculated as a shift in its yield, and the P&L of the liabilities by parallel shifts in the HQM par rate curve. As rates fall, the values of both the liabilities and the asset portfolio increase, as envisioned by the duration hedge. Because of the duration hedge, in fact, the values of both increase by about the same amount for relatively small rate changes. But because the liabilities have a higher convexity, the duration of the liabilities increases more as rates decline. Therefore, for large declines in rates, the increase in value of the liabilities is greater than

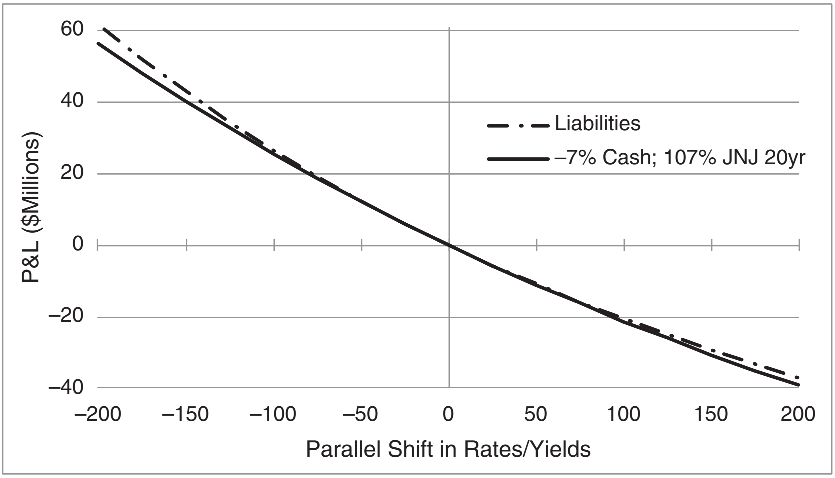

FIGURE 4.11 P&L of the Pension Fund Liabilities in Figure 4.10 and of a Portfolio Borrowing \$9.90 Million and Investing \$149.90 Million in the Johnson & Johnson 2.10s of 09/1/2040. The P&L of the Liabilities Is Computed Under a Parallel Shift of HQM Par Rates. The P&L of the Bond Is Computed by a Shift in Its Yield.

the increase in value of the assets, which leaves the pension fund with a net loss in value.

As rates increase, the values of both the liabilities and the asset portfolio fall, again as envisioned by the duration hedge. But because the liabilities have a higher convexity, the duration of the liabilities falls more as rates increase. Therefore, for large increases in rates, the fall in the value of the liabilities is less than the fall in value of the assets, which, again, leaves the pension fund with a net loss in value.

The pension fund could manage this negative convexity by adjusting its hedge as rates change. As rates fall, and the duration of the liabilities increases more than the duration of the assets, the fund could increase its asset duration by borrowing more cash and buying more bonds. As rates rise, and the duration of the liabilities falls more than that of the assets, the fund could lower its asset duration by selling bonds and paying off some of its borrowing. This solution is fine in theory and sometimes in practice, but it is costly in the sense of requiring potentially frequent portfolio rebalancing.

Another approach is to find an asset portfolio that not only matches the duration of the liabilities, but that also has a convexity equal to or greater than that of the liabilities. There are, of course, many such portfolios that can be formed from cash and the JNJ bonds listed in Table 4.6. For discussion purposes, Table 4.7 lists the portfolio already discussed, along with two additional, two-asset portfolios.

TABLE 4.7 Selected Portfolios of JNJ Bonds Listed in Table 4.6 That Match the Duration of the Liabilities in Figure 4.10, as of Mid-May 2021. Bond Duration and Convexity Are Calculated with Parallel Shifts in Yields. Proportions and Yields Are in Percent.

Asset	Proportion	Asset	Proportion	Wtd-Avg Yield	Duration	Convexity
Cash	−7.05	2.10s of 2040	107.05	2.87	16.66	151.47
0.55s of 2025	24.66	2.25s of 2050	75.34	2.32	16.66	205.47
Cash	30.86	2.45s of 2060	69.14	2.06	16.66	271.84

Each row gives the two assets in the portfolio; the proportion of value in each asset; the weighted-average yield of the portfolio, with weights equal to the proportions of value; the duration of the portfolio; and the convexity of the portfolio. The weighted-average yield of the portfolio is far from a perfect measure of *ex-ante* portfolio return, but will serve for the purposes of this section. Cash, consistent with market levels at the time, is assumed to have a yield of 0.05%. Portfolio duration and convexity equals the weighted average of the duration and convexity of the components, as explained earlier in the chapter.

The first row of the table corresponds to the asset portfolio already analyzed. Its duration equals 16.66, which equals the duration of the liabilities, but its convexity is significantly less than 211, the convexity of the liabilities. The second row gives a portfolio invested in the five-year and 30-year bonds, with the proportions such that the duration of the asset portfolio matches that of the liabilities. The convexity of the portfolio is 205.47, which is quite close to the 211.37 convexity of the liabilities. The third row gives an asset portfolio that once again matches the duration of the portfolio, this time by investing much more in cash and the residual in the 40-year bond. This portfolio has a convexity of 271.84, which exceeds that of the liabilities.

The relative convexities of the asset portfolios in Table 4.7 illustrate the following general principle: among portfolios with the same duration, the more spread out the cash flows of a portfolio, the greater its convexity. All portfolios in the table are constructed to have the same duration, but the first portfolio is essentially all in the 20-year bond. In the asset–liability management context, this portfolio is called a *bullet* portfolio: the liability cash flows, which are spread over many years, are being hedged by a single bond with matching duration. The second asset portfolio in the table is split between five-year and 30-year bonds, while the third is split between cash – essentially a zero-year bond – and 40-year bonds. These two portfolio are called *barbell* portfolios, because they are hedging liabilities using one bond with less duration than the liabilities and one bond with more.

Matched-duration portfolios with cash flows that are more spread out have greater convexities because, as explained earlier, duration is roughly linear in term, while convexity is quadratic. Consider a 10-year zero coupon bond at a yield of zero, which, according to Equations (4.28) and (4.33), has a duration of 10 and a convexity of $10 \times 10.5 = 105$. By contrast, a portfolio with 50% of its value in a five-year zero coupon bond and 50% in a 15-year zero, using the same equations, also has a duration 10, but a convexity of 50% times the convexity of the five-year zero, 22.5, plus 50% times the convexity of the 15-year zero, 232.5, which equals 127.5. Because convexity is quadratic in term, the convexity of the 15-year zero dominates. Hence, the portfolio with the more spread-out cash flows – the one with the longest-term bond – has the greatest convexity.

Given the calculations in Table 4.7, how should the pension fund choose among the alternatives? Figure 4.12 graphs the net P&L of the liabilities and each of the asset portfolios. The net P&L of the liabilities and the bullet portfolio, −7% in cash and 107% in the 20-year bond, is clearly negatively convex: it loses value whether rates fall or rise. The barbell portfolio with 25% in the five-year bond and 75% in the 30-year bond, which matches the duration and approximately matches the duration of the liabilities, gives rise to a net P&L that is very small over a wide range of rates. In other words, a pension fund with this portfolio need not rebalance its asset portfolio unless rates fall significantly. Lastly, the barbell portfolio with 31% in cash and 69% in the 40-year, gives rise to an overall, positively convex position. The combination of assets and liabilities gains value whether rates fall or rise.

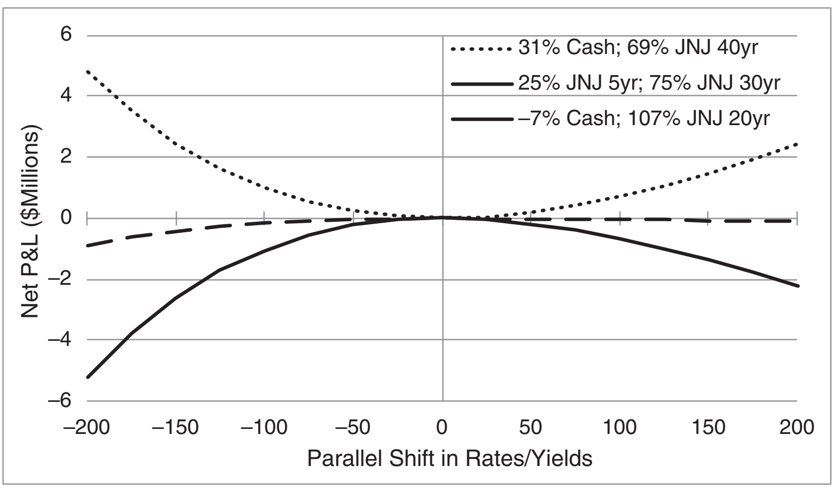

FIGURE 4.12 Net P&L of the Pension Fund Liabilities in Figure 4.10 and the Asset Portfolios in Table 4.7. The P&L of the Liabilities Is Computed Under a Parallel Shift of HQM Par Rates. The P&L of the Bond Is Computed by a Shift in Its Yield.

Does Figure 4.12 prove that the pension fund should always choose the most convex asset portfolio? The answer is no. As mentioned in Section 4.6, evaluating P&L purely as a function of changes in rates does not account for the passage of time. As a rough proxy for return as time passes with rates unchanged, consider the weighted-average yields in Table 4.7, which fall with convexity. In other words, markets recognize the contribution of convexity to P&L as rates change, as shown in Figure 4.12, and, therefore, offer lower yields for highly convex portfolios. As a result, a portfolio with relatively low convexity and relatively high yield underperforms if rates change a lot but outperforms if rates stay about the same. By contrast, a portfolio with relatively high convexity and relatively low yield outperforms if rates change a lot but underperforms if rates stay about the same.

Figure 4.13 illustrates the trade-off between yield and convexity by graphing the yields of the JNJ bonds from Table 4.6 against their durations. The curve has a *concave* shape, meaning that any line connecting two points on the curve is below the curve. Now, in the context of this graph, compare the second portfolio from Table 4.7 – the barbell of the five-year and 30-year bonds – versus the bullet 20-year bond portfolio. The barbell portfolio, with a duration of 16.66 and a weighted-average yield of 2.32%, corresponds to the point on the dotted line at a duration of 16.66. By the concavity of the yield-duration curve, this yield is very much below the yield of the 20-year bond, which has a similar duration.[4] More generally,

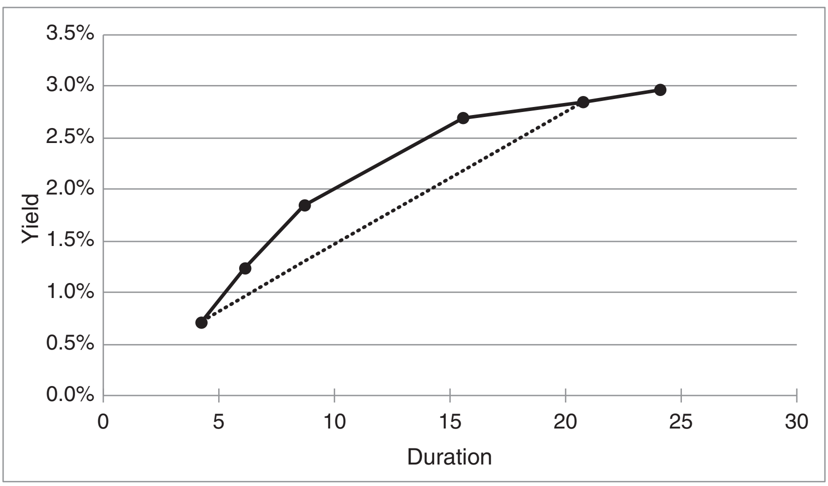

FIGURE 4.13 Yields of the Johnson & Johnson Bonds in Table 4.6 Against Their Durations.

[4]This is a bit imprecise: the duration of the 20-year bond is actually 15.56, while the bullet portfolio has a duration of 16.66.

the concavity of the duration-yield curve implies that the weighted-average yield on portfolios with more spread-out component bonds, which are more convex portfolios, is lower than the yield of bullet portfolios and also of less spread-out – less convex – portfolios. Chapter 8 demonstrates more formally how the shape of the term structure ensures that the benefits of convexity are fairly offset by lower return.

Given the discussion in the last several paragraphs, a pension fund manager who decides to match the duration of assets and liabilities has to decide whether to trade yield for positive convexity. An asset portfolio with a relatively low yield but convexity exceeding that of the liabilities does not have to be rebalanced often to avoid losses, does relatively well when rates move a lot, and does relatively poorly when rates do not move much. An asset portfolio with a relatively high yield but convexity less than that of the liabilities has to be rebalanced often to avoid losses, does relatively poorly when rates move a lot, and does relatively well when rates do not move much. To summarize, then, the choice actually depends on whether rates will change by much or not, that is, on future interest rate volatility!

CHAPTER 5

Key-Rate, Partial, and Forward-Bucket '01s and Durations

The assumption in Chapter 4 that changes in the entire term structure can be described by one interest rate factor may be convenient and appropriate in certain situations, but applications with exposures across the term structure require more realistic foundations. Consider a life insurance company that has liabilities spread out over many years in the future. The insurer could make assumptions about how rates across the term structure typically vary with the 10-year par rate and then hedge all of its liabilities by buying 10-year bonds. But what if, over a particular period, the 30-year rate falls and the 10-year rate stays the same? The present value of liabilities would increase, asset value would be unchanged, and the insurer would experience a net loss in value.

Given that a one-factor framework cannot reliably describe rate changes across the term structure, market participants have pursued three broad strategies. First, build a multi-factor term structure model, based on a combination of data and market analysis. Many active hedge funds and asset managers, who build such models for investment purposes anyway, pursue this path for hedging as well. The Gauss+ model presented in Chapter 9 is an example of this approach. Second, hedge based on empirical analyses of relationships across rates of various terms. Examples of this approach are given in Chapter 6. Third, adopt a robust approach to hedging, which performs well regardless of how the term structure evolves but which leaves little room for subjective views and gives no guidance as to investment decisions. Many asset–liability managers and market makers, who do not see themselves in the business of taking views on term structure movements, pursue this path, which is the subject of this chapter.

The extreme of robust approaches to hedging is *immunization*, that is, hedging away the risk of every single cash flow. An example would be buying eight zero coupon bonds to fund eight semesters of college tuition payments.

No matter what happens to interest rates and the term structure, the maturing zero coupon bonds will fund the assumed payments. Immunization is not practical, however, in many institutional contexts. An insurance company has too many liabilities on too many future dates for each of them to be funded – and therefore hedged – individually. Aside from the operational costs of such a program, the transaction costs of buying relatively small quantities of so many different fixed income securities is likely to be prohibitive.

Practical approaches to robust hedging, therefore, incorporate enough factors to ensure acceptable performance under a wide range of term structure scenarios but limit the number of factors to maintain tractability and contain transaction costs. The three approaches described in this chapter – *key-rate '01s* and durations, *partial '01s* – and *forward-bucket '01s*, have proved popular as striking a sensible balance between hedging effectiveness and practicality or cost. The final section of the chapter briefly connects multi-factor exposures with portfolio volatility.

5.1 KEY RATES: MOTIVATION

The main motivation for multi-factor hedging is to control *curve risk*, that is, the risk that rates across the term structure do not move in parallel or in any fixed relationship. Figure 5.1 illustrates this phenomenon using changes in the term structure of on-the-run US Treasury yields on selected days in June

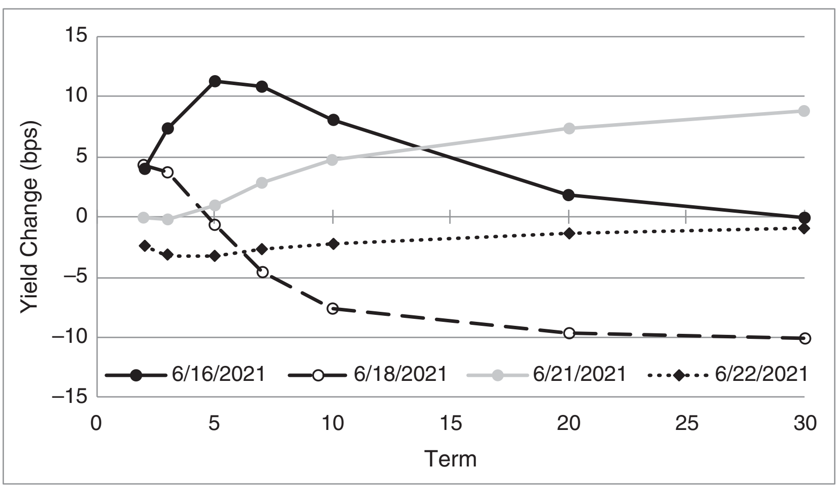

FIGURE 5.1 Changes in the Term Structure of On-the-Run US Treasury Yields, Selected Days, June 2021.

2021. (On-the-run bonds are the most recently issued bonds – and usually the most liquid – in their respective maturity ranges.) Each point on each curve represents a change in an on-the-run Treasury yield over a business day. The 20-year term point on the June 18, 2021, curve, for example, shows that the yield of the on-the-run 20-year Treasury bond fell by 9.7 basis points from June 17, 2021, to June 18, 2021.

The June 16, 2021, curve shows a significant change in the curvature of the term structure: short-term rates rose some; intermediate-term rates rose a lot; and long-term rates rose very little. The June 18, 2021, curve is an example of a *flattening* of the term structure: short-term rates rose, intermediate rates fell some, and long-term rates fell a lot. The June 21, 2021, curve is an example of a *steepening*: short-term rates stayed about the same, while longer-term rates rose significantly. And finally, the June 22, 2021, curve can be very roughly categorized as a parallel shift. From Figure 5.1 as a whole, hedges constructed with the single-factor metrics of Chapter 4 would not have performed reliably over several days in the second half of June 2021.

5.2 KEY RATES: OVERVIEW

Key-rate '01s, or *key-rate DV01s*, along with *key-rate durations*, are designed: i) to describe how the risks of a bond portfolio are distributed along the term structure, and ii) to hedge those risks using some set of bonds, usually the most liquid government bonds of various maturities. To introduce the topic, Table 5.1 shows the key-rate durations for the JPMorgan Government Bond Fund and its benchmark portfolio, the Bloomberg-Barclays US Government Bond Index, as of May 31, 2021. The "Total" durations, in the last row of the table, have similar interpretations as the yield-based durations of Chapter 4. The JPM fund's total duration of 5.73 means that its value increases by approximately 5.73% for a parallel decline in par rates of 100 basis points. The JPM fund, therefore, has less duration risk than its benchmark, which has a duration of 6.78. The other rows of the table give the key-rate durations of these two bond portfolios. From the "5yr" row, for example, the JPM fund increases in value by approximately 0.80% when the five-year par rate declines by 100 basis points, but all other par rates remain unchanged. Similarly, from the 20-year row, the benchmark portfolio increases in value by approximately 1.67% when the 20-year par rate declines by 100 basis points, but all other par rates remain unchanged.

The key-rate duration profile of a portfolio, therefore, decomposes its total duration, that is, its sensitivity to parallel shifts in par rates, into separate durations or sensitivities to changes in individual par rates. Consistent with key-rate durations as a decomposition of total duration, the sum of all

TABLE 5.1 Key-Rate Durations of JPMorgan Government Bond Fund and Its Benchmark Portfolio, the Bloomberg-Barclays US Government Bond Index, as of May 31, 2021.

Key Rate	JPM Government Bond Fund	Bloomberg-Barclays Government Bond Index
6mo	0.03	–
1yr	0.11	0.11
2yr	0.19	0.34
3yr	0.40	0.60
5yr	0.80	0.88
7yr	1.19	0.86
10yr	1.31	0.57
20yr	1.05	1.67
30yr	0.65	1.75
Total	5.73	6.78

key-rate durations of a portfolio is equal or very nearly equal to the total duration of the portfolio.

The key-rate duration profiles in the table say a lot about the exposures of the two portfolios to par rates of different terms. The interest rate risk of the JPM fund is somewhat concentrated in seven- to 20-year par rates, while the risk of the index is concentrated in five- to seven-year and, even more so, in 20- to 30-year rates. Part of the reason for the relative differences in these exposures is that the JPM fund invests in US agency mortgage-backed securities, which are particularly sensitive to seven- to 10-year rates, while the benchmark index contains only government bonds. In any case, as a result of these key-rate duration differences, the two portfolios will perform differently as the shape of the term structure fluctuates.

The JPM fund reports exposures to the nine key rates listed in Table 5.1. The index fund reports only eight key rates, leaving out the six-month rate. The number of key rates chosen involves a trade-off, along the lines of the introduction. With too few key rates, the term structure is likely to move in a way not captured by the framework. With too many, risk management becomes cumbersome and costly.

Along with the number of key rates, the terms of the key rates need be chosen. Because key-rate DV01s and durations are designed to be hedged with a relatively small set of liquid bonds, the most popular choices for key rates are those that correspond to the maturities of on-the-run Treasuries. Including the Treasury's regular issuance of six-month and one-year bills, along with its regular issuance of two-, three-, five-, seven-, 10-, 20-, and 30-year bonds, the JPM and index funds in Table 5.1 have clearly made this popular choice.

5.3 KEY RATES: SHIFTS

The next section applies key-rate metrics and hedging to the problem of the pension fund introduced in Section 4.8. To keep the exposition simple, and to dovetail with the pension fund's long-term liabilities and intention to invest in the Johnson & Johnson (JNJ) bonds listed in Table 4.6, the analysis here is restricted to four key rates with terms of 10, 20, 30, and 40 years.

The key-rate framework imposes no restrictions on how any key rate changes relative to another. But in order to reprice any bond after any shift of key rates, the framework must specify how all par rates move, not just how key rates move. To that end, key-rate analysis makes the following assumptions: i) each key-rate shift leaves all other key rates unchanged; ii) each key-rate shift changes only those par rates that are between adjacent key rates; and iii) changes in par rates due to a key-rate shift decline linearly from the change in that key rate to zero at adjacent key rates. The resulting shifts, in the current example, are shown in Figure 5.2.

The 20-year key-rate shift, depicted by the black, dashed line in the figure, equals one basis point at a term of 20 years. That shift has no effect on par rates with terms less than 10 years – the term of the previous key rate – or on par rates with terms greater than 30 years – the term of the next key rate. Between 10 and 30 years, however, the impact of the shift declines linearly from the 20-year term, which, since the adjacent key rates are both 10 years distant, means a decline of 0.1 basis points per year. For example, the shift increases the 19- and 21-year par rates by 0.9 basis points each; the 18- and 22-year rates by 0.8 basis points each; etc., the 11- and 29-year

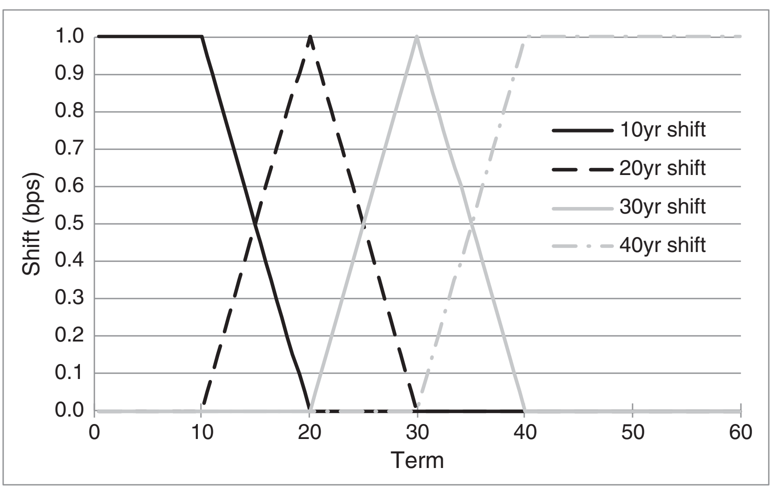

FIGURE 5.2 Key Rate Shifts with Four Key Rates at Terms of 10, 20, 30, and 40 Years.

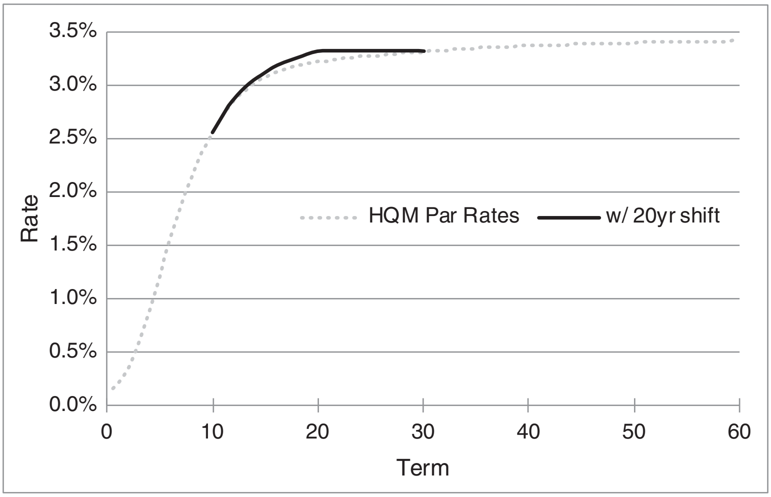

FIGURE 5.3 The HQM Par Rate Curve, as of May 2021, with and Without a 20-Year Key-Rate Shift.

rates by 0.1 basis points each; and the 10- and 30-year par rates, which are the adjacent key rates, by 0.0 basis points each.

When computing the 20-year key-rate DV01 or key-rate duration of a bond or portfolio, all prices are recomputed after shifting the existing par rate curve by the 20-year key-rate shift. Figure 5.3 shows the starting High-Quality Market-Weighted (HQM) corporate par-rate curve as the gray, dotted line, and the 20-year key-rate shift, just described, added to that par-rate curve. The shifted par-rate curve, therefore, is the dotted line out to 10 years, the black line from 10 to 30 years, and the dotted line again from 30 to 60 years. As an aside, while this shifted par-rate curve looks reasonable enough, spot and forward rates implied by this shifted curve may not look as natural or reasonable.

The 30-year key-rate shift in Figure 5.2 is constructed like the 20-year shift. The shift equals one basis point at 30 years and declines linearly on both sides, at 0.1 basis points a year, because the adjacent key rates – at 20 and 40 years – are each 10 years distant. The 10-year and 40-year shifts are constructed somewhat differently, because there is no key-rate at a term less than 10 years or greater than 40 years. The 10-year shift, therefore, equals one basis point for all par rates with terms less than 10 years, and the 40-year shift equals one basis point for all par rates with terms greater than 40 years.

Were all the shifts in Figure 5.2 to occur simultaneously, the result would be a parallel shift of one basis point. This is obvious for par rates of terms less than or equal to 10 years. For terms between 10 and 20 years, which are affected by the 10- and 20-year shifts, note the following. The 10-year shift, which starts at one basis point at 10 years, decreases by 0.1 basis points per year. At the same time, the 20-year shift, which starts at zero at 10 years, increases by 0.1 basis points per year. Hence, the decrease and increase cancel, leaving the shift at one basis point. With similar logic, it is easy to see that simultaneous shifts of all the key rates do result in a parallel shift. This observation explains the sense in which key-rate '01s and key-rate durations decompose the DV01 and duration of Chapter 4: each key-rate exposure is that part of the overall exposure arising from a change in that key rate, holding all other key rates constant.

This section concludes by emphasizing that the assumptions invoked to compute the shifts of par rates between key rates are for convenience rather than to capture market or empirical realities. Most obviously, there is no reason to think that par rates respond linearly to changes in adjacent key rates. More subtly, the assumption that par rates are affected only by adjacent par rates is restrictive as well. If, for example, the 10-year rate falls, while the 20- and 30-year rates stay the same, the 25-year rate might very well change so that the overall shape of the term structure continues to be reasonable. In any case, despite these objections, the formulation of key-rate shifts described here has been widely accepted as a sensible trade-off between reality and tractability.

5.4 KEY RATES: '01S, DURATIONS, AND HEDGING

Unlike their one-factor equivalents in Chapter 4, key-rate DV01s and durations are based on change in price, or percentage change in price, for a change in a single key rate. Mathematically, the DV01 and duration with respect to key rate k are defined as,

$$DV01^k = -\frac{1}{10{,}000}\frac{\partial P}{\partial y^k} \tag{5.1}$$

$$D^k = -\frac{1}{P}\frac{\partial P}{\partial y^k} \tag{5.2}$$

where $\partial P/\partial y^k$ denotes the partial derivative of price with respect to that one key rate. For example, the price of the JNJ 2.10s of 09/1/2040 is 91.2209, at which price its spread to the HQM par-rate curve is −0.514%. After a 20-year key-rate increase of one basis point, the price of the bond – at the same spread – falls to 91.0790. Therefore, the bond's 20-year key-rate '01 is,

$$DV01^{20} = -\frac{1}{10{,}000}\frac{91.0790 - 91.2209}{0.01\%} = 0.1419 \tag{5.3}$$

and its 20-year key-rate duration is,

$$D^{20} = -\frac{1}{91.2209}\frac{91.0790 - 91.2209}{0.01\%} = 15.56 \tag{5.4}$$

TABLE 5.2 Key-Rate DV01s and Hedging of the Pension Liabilities in Figure 4.10 with Par Bonds as of Mid-May 2021.

	Pension Liabilities ($)	2.560s of 05/15 2031	3.215s of 05/15 2041	3.309s of 05/15 2051	3.365s of 05/15 2061
PV/Price	139,786,479	100	100	100	100
10yr Shift	33,192	0.09125	0	0	0
20yr Shift	52,969	0	0.15282	0	0
30yr Shift	57,786	0	0	0.19508	0
40yr Shift	88,766	0	0	0	0.22437
Total	232,713	0.09125	0.15282	0.19508	0.22437
Hedge Face Amount ($millions)		36.375	34.661	29.622	39.562

Table 5.2 takes a preliminary step in using key rates for asset – liability management in the context of the pension fund introduced in Section 4.8. The first column gives the key-rate DV01s of the pension liabilities in Figure 4.10. The total present value of the pension liabilities is about $140 million, and the DV01 of those liabilities is $232,713. The key-rate DV01s decompose that total DV01 into exposures to the four key rates. A 10-year shift of minus one basis point increases the present value of the liabilities by $33,192; a 20-year shift of minus one basis point increases value by $52,969; and so forth. The exposure to the 10-year shift is relatively low because, as evident from Figure 4.10, the liabilities build up gradually over the first 10 years and because the values of cash flows of relatively near terms have relatively low interest rate sensitivities. The exposure to the 40-year key-rate shift is relatively high because, unlike the other shifts, it affects 30 years of cash flows – from 30 to 60 years out – and because the values of those distant cash flows have relatively high interest rate sensitivities.

For expositional purposes, four fictional bonds are created for the rest of the table. (After these results are discussed, the analysis returns to the JNJ bonds.) The maturities of these fictional bonds are chosen to match the terms of the key rates exactly, that is, 10, 20, 30, and 40 years, and their coupons are chosen so that their prices are all par under the HQM par-rate curve in May 2021.

Given the careful selection of these bonds, each has exposure only to the key rate corresponding to its maturity. The 2.560s of 05/15/2031, for example, as of May 15, 2021, are exactly 10-year par bonds. Their price changes only if the 10-year par rate changes. But, by construction, the 20-year, 30-year, and 40-year shifts do not change the 10-year par rate.

Therefore, the 2.560s have exposure only to the 10-year shift, and that exposure also equals the bond's total exposure. With the same logic applying to the other three bonds, key-rate '01s are nonzero only when the term of the key rate equals the maturity of the bond.

Because each of the bonds has such a simple key-rate '01 profile, computing the key-rate hedge of the pension liabilities is correspondingly simple. The objective of the hedge is to purchase an asset portfolio that offsets the key-rate exposures of the liabilities. More specifically, the asset portfolio must have a DV01 exposure to the 10-year shift of \$33,192; to the 20-year shift of \$52,969; etc. But because the only bond that has any exposure to the 10-year shift is the 2.560s of 05/15/2031, the full exposure to that shift has to come from that bond. Mathematically, denoting the face amount of the 10-year bond by F^{10},

$$F^{10} \times \frac{0.09125}{100} = \$33{,}192 \tag{5.5}$$

Solving, F^{10} is about \$36.375 million, which is listed in the bottom row of Table 5.2. Proceeding analogously to hedge the next three key-rate exposures, the equations are,

$$F^{20} \times \frac{0.15282}{100} = \$52{,}969 \tag{5.6}$$

$$F^{30} \times \frac{0.19508}{100} = \$57{,}786 \tag{5.7}$$

$$F^{40} \times \frac{0.22437}{100} = \$88{,}766 \tag{5.8}$$

Each of these equations can be solved on its own to find the hedge face amounts recorded in the table.

To summarize, because the hedging bonds are par bonds with maturities equal to the terms of the key rates, key-rate hedges can essentially be read off the table. The pension fund must purchase \$33,192 in DV01 from 10-year par bonds; \$52,969 in DV01 from 20-year par bonds; and so forth. In this way, the pension fund will be hedged against any combination of key-rate shifts. The four key rates do not need to move according to any model or predetermined pattern. The 10-year key rate might move by +8 basis points, the 20-year by +2 basis points, the 30-year by 0 basis points, and the 40-year by −2 basis points (i.e., something like the May 16, 2021, shift in Figure 5.1), and the hedge should perform relatively well. The caveat, of course, is that the performance can be eroded to the extent that rates between the key rates change in a way markedly different from the shapes of the shifts in Figure 5.2. This caveat, however, is more of an issue in this simplified example, with only four key rates over a maturity range of

TABLE 5.3 Key-Rate DV01s and Hedging of the Pension Liabilities in Figure 4.10 with Johnson & Johnson Bonds. Indicative Prices Are as of Mid-May 2021.

	Pension Liabilities ($)	1.30s of 09/01 2030	2.10s of 09/01 2040	2.25s of 09/01 2050	2.45s of 09/01 2060
PV / Price	139,786,479	95.357	91.221	88.154	88.157
HQM Sprd (bps)	0	−56.4	−51.4	−45.4	−39.2
10yr Shift	33,192	0.08577	0.00597	−0.00393	−0.00250
20yr Shift	52,969	0	0.14191	0.00707	−0.00473
30yr Shift	57,786	0	0	0.18494	0.00871
40yr Shift	88,766	0	0	0	0.21548
Total	232,713	0.08577	0.14788	0.18808	0.21695
Hedge Face Amount ($millions)		38.647	37.239	29.306	41.195

60 years, than it is in practice, for example, with the eight or nine key rates over a maturity range of 30 years used by the JPM Government Bond Fund and the Bloomberg-Barclays US Government Bond Index in Table 5.1.

With the basics of key-rate hedging understood, the fictional par bonds created for Table 5.2 are replaced by the approximately 10-year, 20-year, 30-year, and 40-year JNJ bonds listed in Table 4.6. The updated analysis is presented in Table 5.3. The row "HQM Sprd" gives the spread of each of the JNJ bonds to the HQM par-rate curve as of mid-May 2021. These spreads are all negative, because, with its triple-A rating, JNJ bonds sell at lower rates than HQM par-rates, which represent double- and single-A credits. In any case, the bond spreads in the table are kept constant as key-rate shifts are applied. No spread is attached to the pension fund liabilities, because, by assumption, their present value is computed directly from the HQM par-rate curve.

A major difference between Table 5.2, which features fictional, par bonds, and Table 5.3, which features the JNJ bonds, is that the key-rate DV01 profile of the JNJ bonds is more complex. As of mid-May 2021, the 1.30s of 09/1/2030 mature in less than 10 years, which means that only the 10-year key-rate shift affects their valuation. Their key-rate profile, therefore, is completely concentrated in the risk of a 10-year shift. The other bonds, however, which do not mature in exactly 20-, 30-, and 40-years and which are not priced at par, have small positive or negative exposures to key rates other than those corresponding most closely to their maturities.

Because the bulk of each bond's exposure is still at the key rate closest to its maturity, the hedging problem in Table 5.3 is not conceptually different from the earlier table. Solving for the hedging portfolio is more tedious, however, because no key-rate exposure of the liability can be handled in isolation. All of the bonds have exposures to the 10-year shift; three have exposures to the 20-year shift; and so forth. Consequently, the following equations (with key-rate '01s rounded for readability) need to be solved together,

$$F^{10}\frac{0.086}{100} + F^{20}\frac{0.006}{100} + F^{30}\frac{-0.004}{100} + F^{40}\frac{-0.003}{100} = \$33{,}192 \qquad (5.9)$$

$$F^{20}\frac{0.142}{100} + F^{30}\frac{0.007}{100} + F^{40}\frac{-0.005}{100} = \$52{,}969 \qquad (5.10)$$

$$F^{30}\frac{0.185}{100} + F^{40}\frac{0.009}{100} = \$57{,}786 \qquad (5.11)$$

$$F^{40}\frac{0.215}{100} = \$88{,}766 \qquad (5.12)$$

Equation (5.9) says that the combined 10-year key-rate '01 of the asset portfolio must offset the 10-year key-rate '01 of the liabilities. Equation (5.10) says the same for the 20-year key rate; Equation (5.11) for the 30-year key rate; and Equation (5.12) for the 40-year key rate. The solution, given in the bottom row of Table 5.3, is not the same as that in Table 5.3 but is certainly of the same order of magnitude.

5.5 PARTIAL '01S AND PV01

Interest rate swap trading desks are firmly in the category of market participants that require robust, multi-factor frameworks for hedging. These desks are mostly market makers in swaps, regularly paying fixed to and receiving fixed from customers, while trying to hedge any residual interest rate and curve risk. In order to value these trading books, swap rate curves are constructed to fit or pass through rates of all terms that trade with significant liquidity. There are typically quite a few of these fitting rates or fitting points, perhaps between 15 to 25, depending on the application and currency. In any case, swap desks often leverage their curve-building machinery to compute *PV01s* (i.e., *present value of an '01*) and *partial '01s* or *partial PV01s*. The term PV01 is standard in swap markets, but it is really a synonym for the general conception of DV01 described in the early sections of Chapter 4. The term is likely preferred so as to differentiate between that general conception and yield-based DV01, which is particularly widespread in bond markets.[1]

[1]While this section focuses on swaps trading desks, any business that trades heavily in swaps, whether to take positions or to hedge, may choose to use partial '01s as well. One example would be a hedge fund that specializes in relative value trades

The PV01 of a portfolio of swaps is found as follows: compute its present value with the current swap rate curve; shift all fitted rates by one basis point; refit the curve; recompute the present value; and subtract one present value from the other. The sign convention for PV01 is the same as for DV01: positive numbers mean that value increases when rates fall.

A partial '01 of a portfolio is defined for each rate used to fit the swap curve. To find a partial '01 with respect to a particular fitted rate, compute the present value of the portfolio with the current swap curve; shift that fitted rate, keeping all other fitted rates the same; refit the curve; recompute the present value; and subtract one present value from the other, again observing the usual sign convention.

The advantages and caveats of using key-rate '01s apply to partial PV01s as well. Partial '01s decompose the overall PV01 into exposures to the individual fitted swap rates. Hedging is simple by construction: given a portfolio's partial '01 with respect to a particular fitted rate, offset that exposure by receiving or paying fixed in a swap whose term corresponds to that fitted rate. In this respect, hedging is actually simpler with partial '01s than with key-rate '01s. Because new, hedging swaps are always at par (see Chapter 2), each partial '01 can be hedged individually, as in the simplified example of Table 5.2, where key-rate exposures are hedged with par bonds.

With respect to caveats, refitting a curve after shifting one fitted swap rate, keeping all other fitted rates the same, can result in oddly shaped term structures of swap rates and even more so, of spot and forward rates. Furthermore, additional problems can arise if the curve-fitting methodology is not local enough, that is, for example, if shifting the 10-year swap rate, keeping other rates the same, unintentionally moves rates between the fitted five- and seven-year rates. In that case, the present value of a six-year swap would show an unintended sensitivity to the 10-year swap rate, and, even though five- and seven-year swap rates are included as hedging instruments, the methodology would call for hedging six-year swaps not only with five- and seven-year swaps, but also with some amount of 10-year swaps. This outcome is fine if intended, but not fine if just an artifact of the curve-fitting methodology. Because the success of using partial '01s is very dependent on the properties of the curve fitting methodology, many market participants are careful to use robust curve fitting methodologies or use forward-bucket '01s, which are described in the next section.[2]

Before concluding this section, it is noted that the term *CV01* denotes the change in value of a swap for a one-basis-point decrease in its coupon or

---

across the swaps curve. Another would be a credit hedge fund or asset manager that hedges away some or all of the interest rate risk of bond positions with swaps. In that case, it is typically convenient to use the swap curve for all discounting and interest rate risk metrics and to handle the idiosyncrasies of individual issuers and bond issues with spreads against the swap curve.

[2]There is a brief discussion of robustness in curve fitting in Section 12.3.

fixed swap rate. CV01 is particularly easy to compute, because a moment's reflection reveals that it is proportional to the sum of the discount factors used to discount the fixed payments of the swap. CV01 and PV01 are quite different conceptually, with the former shifting the coupon and the latter shifting discount rates. The two are sometimes used interchangeably and confused, however, perhaps because they are essentially equal for par swaps, which constitute the bulk of swaps trading activity.[3]

5.6 FORWARD-BUCKET '01S

Key-rate and partial '01s are particularly convenient in that they translate directly into hedging with liquid instruments, whether on-the-run or near on-the-run Treasuries or swaps of the most liquid terms. But the shifts themselves are not very intuitive. Shifting the 10-year par rate by a basis point does not mean that the 10-year segment of the term structure shifts, while all other segments stay constant: shifting the 10-year par rate has implications for discounting all cash flows out to 10 years. Furthermore, as mentioned earlier, shifting one par rate, while holding all other rates constant, can imply odd shifts in spot or forward rates. *Forward-bucket '01s*, by contrast, are more easily interpretable in terms of changes in the shape of the term structure. They do not, however, provide the same direct visibility into hedging with liquid instruments as do key-rate and partial '01s.

This section illustrates the use of forward-bucket '01s with reference to the general obligation bonds of the Town of Wellesley, Massachusetts, issued in May 2020.[4] Table 5.4 lists the first nine bond series sold at that time. Note that, for the tax reasons explained in Section O.2, it is common for municipalities to issue bonds with a coupon of 5% even though market yields are much lower.

Using the yields in Table 5.4, the analytical tools of Part One can be used to calculate a term structure of forward rates, which is graphed in Figure 5.4.[5] The idea behind forward-bucket '01s is to shift "buckets," that is, individual segments of the forward term structure, directly. For purposes of exposition, this section breaks the term structure into three segments: one to two years; three to five years; and six to nine years. Figure 5.4 shows,

[3]For a rough understanding of this equality, note from Equation (4.21) that, with a curve flat at the rate y, the change in the price of a bond for a one-basis-point change in coupon (e.g., from a dollar coupon of 2 to 2.01), is $(1/100y)(1-(1+y/2)^{-2T}$, which is exactly the same as the DV01 of a par bond in Equation (4.29).

[4]General obligation bonds are discussed in Chapter O.

[5]For simplicity, it is assumed that bonds pay annually rather than semiannually. Absent this assumption, other assumptions would have to be made about discount factors on September 1 of each year.

TABLE 5.4 Selected General Obligation Refunding Bonds Issued by the Town of Wellesley, Massachusetts, in May 2020. Coupons and Yields Are in Percent.

Maturity	Principal	Coupon	Yield
06/01/2021	1,175,000	5.00	0.75
06/01/2022	1,205,000	5.00	0.80
06/01/2023	1,210,000	5.00	0.84
06/01/2024	1,225,000	5.00	0.90
06/01/2025	1,235,000	5.00	0.96
06/01/2026	1,250,000	5.00	1.01
06/01/2027	1,255,000	5.00	1.06
06/01/2028	1,265,000	5.00	1.11
06/01/2029	1,270,000	5.00	1.18

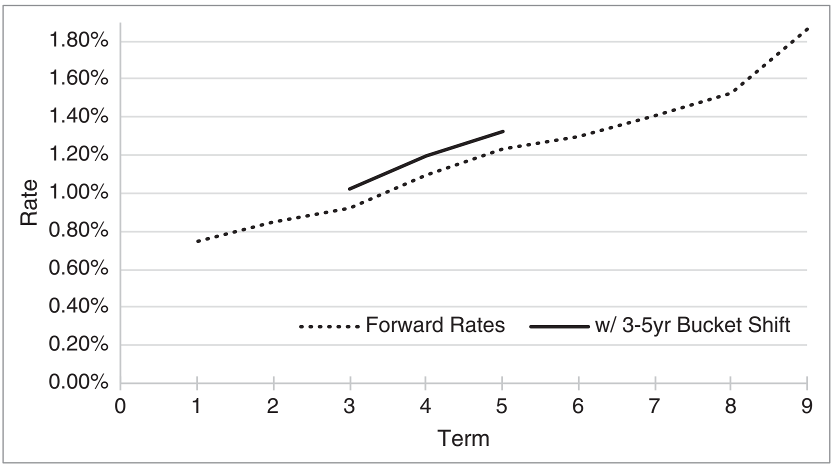

FIGURE 5.4 Town of Wellesley General Obligation Forward Rate Curve with a 10-Basis-Point Shift of the Three- to Five-Year Forward Bucket, as of June 1, 2020.

for example, a +10-basis-point shift of the three- to five-year segment. In practice, practitioners choose buckets most suitable to their businesses. A money-market desk, for example, which trades very heavily in short-term instruments, might have many buckets in the short end of the curve and few further on. A swaps desk, on the other hand, which trades evenly across maturities from one to 30 years, might not have nearly as many buckets in the short end but many more evenly spaced buckets across the rest of the curve.

Forward-bucket ’01s are computed analogously to any other interest rate risk metric: price the bond or portfolio with the current term structure; shift a forward-bucket by one basis point; reprice the bond or

TABLE 5.5 Forward-Bucket Exposures and Hedging of Selected Town of Wellesley General Obligation Refunding Bonds, as of June 1, 2020. The 5s of 06/01/2029 are Hedged with the 1.205s of 06/01/2029.

(1)	(2) 5s of 06/01 2022	(3) 5s of 06/01 2025	(4) 5s of 06/01 2029	(5) 1.205s of 06/01 2029	(6) Hedged 5s of 06/01 2029
1-2yr Shift	0.02099	0.02324	0.02578	0.01972	0.00239
3-5yr Shift	0	0.03112	0.03492	0.02862	0.00098
6-9yr Shift	0	0	0.03983	0.03642	−0.00337
Total	0.02099	0.05437	0.10053	0.08475	0

portfolio with the new term structure; and subtract one price from another, observing the usual sign convention. Table 5.5 reports forward-bucket '01 profiles computed along these lines. The second to fourth columns give the forward-bucket '01s for three of the Wellesley bonds, a two-year bond, a five-year bond, and a nine-year bond, all as of June 1, 2020. For expositional purposes, the fifth column gives the forward-bucket '01s of a fictional nine-year par bond, discounted with the same curve. The final row of the table sums all the forward-bucket '01s, which gives the increase in price after shifting all forward rates down by one basis point. Hence, the forward-bucket '01s in the table decompose these totals into exposures to individual segments of the term structure.

The two-year bond, which has no cash flows past two years, has no exposure to the three- to five-year and the six- to nine-year forward-bucket shifts. Similarly, the five-year bond has no exposure to the six- to nine-year shift. As intended, forward-bucket '01s clearly convey exposures to individual segments of the term structure. For example, if forward rates in the three- to five-year segment of the curve fall by one basis point, the price of the 5s of 06/1/2029 increases by 3.5 cents per 100 face amount.

The usefulness of forward-bucket '01s to assess curve risk is illustrated in the sixth column of Table 5.5. Say that a market maker hedged the 5s of 06/01/2029 with the hypothetically liquid, fictional 1.205s of 06/01/2029 using the parallel shift, that is, the total '01s in the last row of the table. Against a purchase of 100 face amount of the 5s of 06/01/2029, the market maker would sell $100 \times 0.10053/0.08475 = 118.62$ of the 1.205s. By definition, the total '01 of this hedged portfolio is zero, but it does have curve risk, which is computed in the last column of the table.

The exposure of the hedged portfolio to each forward bucket is the sum of the exposures of the two bond positions. For the one- to two-year forward bucket, for example, the portfolio exposure is,

$$100\frac{0.02578}{100} - 118.62\frac{0.01972}{100} = 0.00239 \tag{5.13}$$

The other two bucket exposures are computed similarly. Taken as a whole, the forward-bucket '01 profile of the hedged portfolio, with the first two exposures positive and the third negative, shows that the market maker's position loses money if the forward curve flattens. For example, in a particular flattening scenario, if one- to five-year forward rates increase, while six- to nine-year forward rates decrease, the position will lose money due to each of the three forward rate shifts. With curve risk quantified in this way, the market maker can choose between hedging out all curve risk, at some cost, or bearing the risk until the position can be unwound.

Hedging a bond or portfolio with forward-bucket '01s follows along the same lines as hedging with key-rate '01s. If there are five forward buckets, for example, then five hedging instruments are required. The five unknowns are the face amount of the five hedging instruments, and each of five equations sets the net '01 of its corresponding forward bucket to zero. While this procedure is perfectly straightforward, an approximate solution is not as immediately apparent as in the case of key-rate '01s. Inspection of the sixth column of Table 5.5 does not reveal the approximate position of a bond or portfolio of bonds that hedges out the residual curve risk. By contrast, inspection of the dollar key-rate DV01s in the second column of Table 5.3 readily reveals the approximate trades required in the hedging bonds.

5.7 MULTI-FACTOR EXPOSURES AND PORTFOLIO VOLATILITY

As discussed in this chapter and Chapter 4, risk metrics are useful for quantifying the effect of a change in rates on prices and for hedging interest rate risk. This section makes the point that risk metrics are also useful for quantifying portfolio volatility.

Say that a portfolio has a DV01 of $10,000 and that interest rates have a volatility of 100 basis points per year. It follows that the portfolio has an annual volatility of $10,000 times 100, or $1 million. This, of course, is a one-factor formulation of interest rate volatility.

The motivation for multi-factor exposures, however, is that portfolios are exposed to changes in interest rates all along the curve and that these changes are not perfectly correlated. The discussion here quantifies portfolio volatility in the simple, multi-factor setting of two key rates. The ideas, however, are easily extended to additional key rates and to partial and forward-bucket '01s.

To estimate portfolio volatility with two key rates, follow these steps. First, empirically estimate the volatility of changes in each key rate and the correlation between them. Second, compute the key-rate '01s of the portfolio. Third, compute the variance or volatility of the portfolio. To illustrate this third step, denote the change in the value of the portfolio by ΔP; the change in the two key rates by Δy_1 and Δy_2; and the key rates of the portfolio by $KR01_1$ and $KR01_2$. Then, by the definition of key-rate '01s,

$$\Delta P \approx KR01_1 \times \Delta y_1 + KR01_2 \times \Delta y_2 \tag{5.14}$$

Furthermore, denoting the volatility of changes in portfolio value and changes in the two key rates by σ_P, σ_1, and σ_2, respectively, and the correlation between changes in the key rates by ρ, Equation (5.14) and the rules for computing variances and standard deviations imply that,

$$\sigma_P \approx \sqrt{KR01_1^2\sigma_1^2 + KR01_2^2\sigma_2^2 + 2KR01_1 KR01_2 \rho\sigma_1\sigma_2} \tag{5.15}$$

Of course, as the number of key rates or, more generally, factors, increases, more volatility and correlation inputs are required.

CHAPTER 6

Regression Hedging and Principal Component Analysis

The risk metrics and hedges of Chapters 4 and 5 assume relationships across rates of different terms. The assumptions are typically motivated by a combination of economic theory, empirical analysis of historical data, and views about future economic and financial developments. This chapter introduces approaches that rely explicitly on historical data. It would be an oversimplification, however, to categorize the techniques of Chapters 4 and 5 as subjective, while categorizing the approaches of this chapter as objective. Empirical methods also require assumptions, such as the number of rates or instruments used in the analysis, the particular rates or instruments chosen, and the historical time period selected for estimation.

The first section of this chapter describes single-variable regression hedging in the context of hedging a 40-year Johnson & Johnson (JNJ) bond with a 30-year Treasury. The second section describes two-variable regression hedging in the context of a relative value trade of 20-year versus 10- and 30-year Treasuries. The third and fourth sections discuss two other issues that arise in the context of regression hedging, namely, the choice between *level* and *change* regressions, and *reverse* regressions.

One conceptual problem with using regression hedging in practice is that each regression is essentially a different model of the term structure of interest rates with different underlying assumptions. Consider, for example, the manager of a trading desk in which some traders are using single-variable regressions and some two-variable regressions, or some are estimating regressions from one month of historical data and others from six months.

The final section of the chapter introduces a unified empirical description of how the entire term structure evolves, namely, *principal component analysis (PCA)*. PCA provides both an empirical hedging methodology, which can be used consistently across the term structure, along with easily interpreted descriptions of how the term structure fluctuates over time. While presentations of PCA tend to be highly mathematical, great effort has been made in this chapter to make the material more broadly accessible.

6.1 SINGLE-VARIABLE REGRESSION HEDGING

This section considers the problem of a market maker or a relative value trader on May 14, 2021, who purchases $100 million face amount of the JNJ 2.450s of 09/01/2060 and hedges the resulting interest rate risk by selling the US Treasury 1.625s of 11/15/2050. Because there is no 40-year Treasury bond outstanding, the 1.625s of 11/15/2050, with about 30 years to maturity, are selected as the best alternative. Table 6.1 gives the coupons, maturity dates, yields, and DV01s of these two bonds, along with those of two other bonds that are referenced in the next section.

The trader can choose the face amount of the Treasury bond in the hedge using the ratio of DV01s, along the lines of Chapter 4. In this case, the trader sells $100 million × 0.2124/0.1910, or $111.2 million. As discussed in Chapter 4, however, this hedge assumes that the yields of the JNJ and Treasury bonds move up or down in parallel. But because the JNJ bonds sell at a changing corporate spread to Treasuries, and because 40- and 30-year rates are not perfectly correlated, as emphasized in Chapter 5, there is good reason to question the assumption of parallel yield shifts in this case.

The scatter plot in Figure 6.1 shows the daily changes in yield of the JNJ bonds against those of the Treasury bond from January 19, 2021, to May 14, 2021, which is a window of about four months before the date of the hedge. The line in the figure is the regression line fitted through the data, which is discussed presently. The figure teaches two lessons. First, there is a lot of variation in the relationship between these changes. Some days, the Treasury yield changes by more (e.g., the point on the graph at (−18.2, −11.1)); some days by less (e.g., the point (−3.3, −7.7)); and some days even in opposite directions (e.g., the point (2.0, −2.2)). Second, from the slope of the line, the average relationship between yield changes is less than one-to-one; that is, the change in the yield of the JNJ bonds tends to be less than the change in the yield of the Treasury bonds.

Figure 6.1 implicitly assumes that the most relevant period for designing an empirical hedge is the recent, immediate past. This assumption is often reasonable, but there are times and situations in which some earlier period seems more relevant. For example, if the Federal Reserve is expected to raise

TABLE 6.1 Yields and Yield-Based DV01s for the JNJ 2.450s of 09/01/2060 and Selected US Treasury Bonds, as of May 14, 2021. Yields Are in Percent.

Issuer	Bond	Yield	DV01
JNJ	2.450s of 09/01/2060	2.962	0.2124
Treasury	0.875s of 11/15/2030	1.601	0.0847
Treasury	1.375s of 11/15/2040	2.246	0.1446
Treasury	1.625s of 11/15/2050	2.364	0.1910

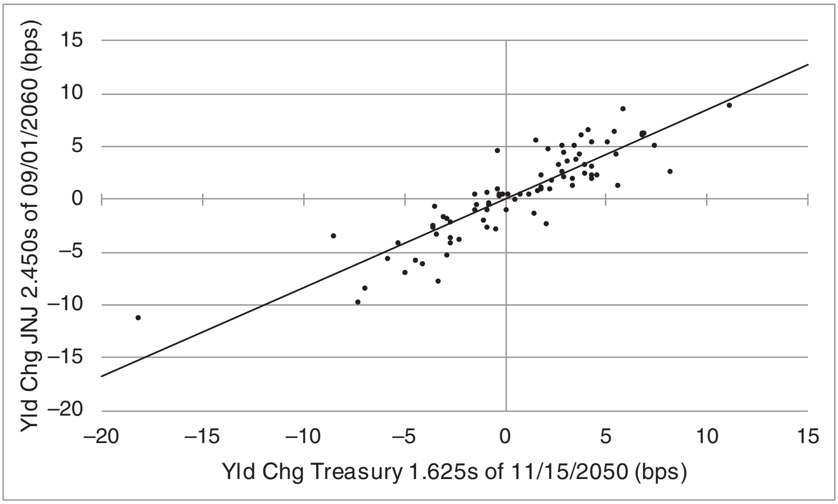

FIGURE 6.1 Regression of Daily Changes in Yields of the JNJ 2.450s of 09/01/2060 on Daily Changes in Yields of the Treasury 1.625s of 11/15/2050, from January 19, 2021, to May 14, 2021.

short-term rates in the immediate future, a similar episode in the past might be more relevant to the future than the recent past, in which the Federal Reserve left rates unchanged or lowered them. Choosing the length of the observation or estimation window is also part of the art of regression hedging. Too short a window might fail to furnish statistically reliable estimates, but too long a window might include less relevant historical data.

In light of the empirical evidence in Figure 6.1, the trader might very well choose to: i) adjust the hedge ratio to account for the less than one-to-one relationships between changes in yields; and ii) measure the variation around the average relationship to gain a better understanding of the risk of the hedged position. Regression analysis is a tool with which to achieve both of these objectives.

Let Δy_t^{JNJ} and Δy_t^{30} be the changes in yields of the JNJ and 30-year Treasury bonds on date t, respectively. A regression model linking these changes is,

$$\Delta y_t^{JNJ} = \alpha + \beta \Delta y_t^{30} + \epsilon_t \tag{6.1}$$

Equation 6.1 says that the *dependent variable*, here the change in the yield of the JNJ bond, equals: a constant or intercept, α; plus a slope, β, times the *independent* variable, here the change in the yield of the 30-year Treasury bond; plus an error term, ϵ_t. The unknown constant and slope parameters are estimated from the data, in a manner explained presently. These estimated parameters, denoted $\hat{\alpha}$ and $\hat{\beta}$, respectively, can then be used for prediction. Given the change in the Treasury bond yield on date t, the predicted change in the yield of the JNJ bonds on that date, denoted $\Delta\hat{y}_t^{JNJ}$, is,

$$\Delta\hat{y}_t^{JNJ} = \hat{\alpha} + \hat{\beta}\Delta y_t^{30} \tag{6.2}$$

and the realized error or *residual* on that day, $\hat{\epsilon}_t$, is given by,

$$\hat{\epsilon}_t = \Delta y_t^{JNJ} - \hat{\alpha} - \hat{\beta}\Delta y_t^{30} \tag{6.3}$$

$$= \Delta y_t^{JNJ} - \Delta\hat{y}_t^{JNJ} \tag{6.4}$$

For example, say that the estimated constant and slope parameters are 0 and 0.84, respectively, and that the Treasury bond yield changes by −18.2 basis points. Then, by Equation (6.2), the predicted change in the yield of the JNJ bond is $0 + 0.84 \times (-18.2) = -15.3$ basis points. If, furthermore, the actual change in the JNJ bond is −11.1 basis points, then, by (6.3) or (6.4), the realized error or residual is $-11.1 - (-15.3)$ or 4.2 basis points. In Figure 6.1, this residual can be thought of as a vertical line dropped from the data point, (−18.2, −11.1), to the regression line.

Least-squares estimation of the unknown parameters finds $\hat{\alpha}$ and $\hat{\beta}$ to minimize the sum of the squares of the residuals over the observation period,

$$\sum_t \hat{\epsilon}_t^2 = \sum_t \left(\Delta y_t^{JNJ} - \hat{\alpha} - \hat{\beta}\Delta y_t^{30}\right)^2 \tag{6.5}$$

where the equality follows from Equation (6.3). Squaring of the errors ensures that offsetting positive and negative errors are not considered as acceptable as zero errors, and that large errors in absolute value are penalized heavily relative to smaller errors.

Least-squares estimation assumes that the regression model is a true description of the dynamics of the dependent and independent variables, that the errors across time have the same probability distribution, that they are independent of each other, and that they are uncorrelated with the independent variable. Under these assumptions, least-squares parameter estimates are linear, unbiased, consistent, and efficient.[1]

Least-squares estimation is available in many statistical packages and spreadsheets. Table 6.2 gives a typical summary output from estimating Equation (6.1) using the data shown in Figure 6.1. The estimate of the slope

[1]A linear estimator is linear in the observations of the dependent variable. The expectation of an unbiased estimator of a parameter equals the true value of that parameter. A consistent estimator of a parameter, with enough data, becomes arbitrarily close to the true value of the parameter. And an efficient estimator has the minimum possible variance among linear estimators.

coefficient, $\hat{\beta}$, is 0.842, which says that, on average, the change in the yield of the JNJ bond is only 0.842 times the change in the yield of the Treasury bond, which is very different from a parallel shift. The estimate of the constant, $\hat{\alpha}$ is not very different from zero, which is typically the case in regressions of this sort. From an economic perspective, it would be odd if, over an extended period of time, changes in the yield of the JNJ bond tended to be positive or negative when there is no change in the yield of the Treasury bond. The line in Figure 6.1 is the *fitted regression line*, which is Equation (6.2) with its estimated coefficients,

$$\Delta\hat{y}_t^{JNJ} = 0.060 + 0.842\Delta y_t^{30} \tag{6.6}$$

Table 6.2 also gives the standard errors of the constant and slope coefficients, which provide confidence intervals around the estimates: the interval of each estimate plus or minus two standard errors falls around the true parameter values approximately 95% of the time. In this regression, the confidence intervals are 0.060 plus or minus 2 times 0.223, or $(-0.386, 0.446)$, and 0.842 plus or minus 2 times 0.051, or $(0.740, 0.944)$. Hence, because the confidence interval around the estimated constant includes zero, the hypothesis that $\alpha = 0$ cannot be rejected with 95% confidence. But, because the confidence interval for the slope coefficient does not include one, the hypothesis that $\beta = 1$ can be rejected with 95% confidence. Hence, the hypothesis of parallel shifts in the yields of the two bonds is rejected by the data.

Table 6.2 reports that the R-squared of the regression is 77.5%, meaning that 77.5% of the variance of changes in the JNJ yield can be explained by the model, that is, by changes in the yield of the 30-year Treasury bond. In a one-variable regression, the R-squared is just the square of the correlation between the dependent and independent variables, which gives a correlation here of $\sqrt{77.5\%}$, or 88.0%. That these statistics are well below 1.0 indicates that hedging in this case does not come close to eliminating all interest rate risk.

TABLE 6.2 Regression of Daily Changes in Yields of the JNJ 2.450s of 09/01/2060 on Daily Changes in Yields of the Treasury 1.625s of 11/15/2050, from January 29, 2021, to May 14, 2021.

No. of Observations	**82**	
R-Squared	**77.5%**	
Standard Error	**2.00**	
Regression Coefficients	**Value**	**Std. Error**
Constant ($\hat{\alpha}$)	0.060	0.223
Chg 30yr Treasury Yield ($\hat{\beta}$)	0.842	0.051

The remaining statistic to be discussed in Table 6.2 is the standard error of the regression, which is essentially the standard deviation of the realized errors or residuals, as defined in Equations (6.3) and (6.4).[2] This standard error is a measure of how well the model fits the data and is in the same units as the dependent variable, in this case, basis points. Roughly speaking, then, the standard deviation of the errors in explaining daily changes in the yield of the JNJ bond with daily changes in the yield of the Treasury bond is two basis points. This statistic is particularly useful in describing the risk of a regression-based hedge, as discussed presently.

All the results in Table 6.2 are *in-sample*; that is, they are computed from the particular data sample used to estimate the regression model. Relying on these results for hedging assumes that the future will be sufficiently like this particular historical period. The success or failure of this assumption is discussed at the end of the section.

Turning now to regression hedging, assume for the moment that the yield of the JNJ bonds changes by exactly $\hat{\beta}$ basis points for every one-basis-point change in the yield of the Treasury bonds. Let F^{JNJ}, $DV01^{JNJ}$, F^{30}, and $DV01^{30}$ be the face amounts and DV01s of the JNJ and 30-year Treasury bonds, respectively. Then, the position is hedged against changes in yields if,

$$F^{JNJ}\frac{DV01^{JNJ}}{100}\hat{\beta} + F^{30}\frac{DV01^{30}}{100} = 0 \tag{6.7}$$

$$F^{30} = -F^{JNJ}\frac{DV01^{JNJ}}{DV01^{30}}\hat{\beta} \tag{6.8}$$

Plugging in numbers, \$100 million for the face amount of the JNJ bonds to be hedged, DV01s from Table 6.1, and $\hat{\beta}$ from Table 6.2,

$$F^{30} = -\$100mm\frac{0.2124}{0.1910}0.842 = -\$93.6mm \tag{6.9}$$

The yield-based DV01 hedge for the JNJ bonds, which is given by Equation (6.9), is given earlier without the slope coefficient of 0.842 as \$111.2 million. The regression hedge of (6.9) sells only \$93.6 million because it recognizes that the yield of the JNJ bonds does not move as much as the yield of the 30-year Treasury bond. Hence, fewer Treasury bonds need be sold to hedge the interest rate risk of the JNJ bonds.

[2]A standard error is actually the sum of the squared residuals divided by the number of observations minus two, while a standard deviation divides by the number of observations minus one. Note than, in a regression with a constant, the average of the residuals is zero by construction.

Regression hedges are sometimes described in terms of *risk weights*. Rearranging terms in Equation (6.7) or (6.8),

$$\frac{-F^{30} \times DV01^{30}/100}{F^{JNJ} \times DV01^{JNJ}/100} = \hat{\beta} = 84.2\% \tag{6.10}$$

In words, the left-hand side of the equation is the DV01 of the hedge as a fraction of the DV01 of the bonds being hedged. The risk weight of a yield-based DV01 hedge is always 100% – the DV01s of the two sides of the trades are, by construction, equal. In this regression hedge, however, the DV01 of the Treasury bonds is only 84.2% of the DV01 of the JNJ bonds. In general, as Equation (6.10) shows, the risk weight of a regression hedge exactly equals the estimated slope coefficient, $\hat{\beta}$.

The best argument for the regression hedge is actually not the earlier assumption that bond yields change exactly according to the regression model. Write the P&L of the position as,

$$P\&L = -F^{JNJ}\frac{DV01^{JNJ}}{100}\Delta y_t^{JNJ} - F^{30}\frac{DV01^{30}}{100}\Delta y_t^{30} \tag{6.11}$$

where the negative signs reflect that a positive face amount with a positive change (i.e., increase) in yield lowers P&L. It can then be shown that the regression hedge in Equation (6.8) minimizes the variance (6.11). (See Appendix A6.1.) In other words, to the extent that P&L variance is an appropriate measure of risk, the regression hedge minimizes the risk of the hedged position.

Appendix A6.1 also derives the standard deviation of the regression-hedged P&L. Denote this standard deviation by $\sigma_{P\&L}$ and the standard deviation of the residuals by σ_ϵ. Then,

$$\sigma_{P\&L} = \left| F^{JNJ}\frac{DV01^{JNJ}}{100} \right| \sigma_\epsilon \tag{6.12}$$

where $|\cdot|$ is the symbol for absolute value, so that the standard deviation is positive whether the original position is long (positive F^{JNJ}) or short (negative F^{JNJ}). In words, the P&L of the hedged position is the DV01 of the position being hedged times the standard error of the regression residuals. Intuitively, the hedged P&L on any given day is exactly zero if the yield of the JNJ bonds moves by 0.842 basis points times the change in the Treasury yield. But if the residual is one basis point, so that the yield of the JNJ bond increases by one basis point more than that, the hedged position loses the DV01 of the JNJ bonds; and if the residual is minus two basis points, then the hedged position gains twice the DV01 of the JNJ bonds; etc. Hence, the volatility of the hedge is proportional to the variability of the residuals.

Applying Equation (6.12) to the case at hand, the DV01 of the JNJ bond position is \$100 million × 0.2124/100, or \$212,400, and the standard error of the regression, reported in Table 6.2, is two basis points per day. Therefore, the standard deviation of the hedged P&L in the sample is \$212,400 × 2 or \$424,800 per day. Whether this is too much risk or not depends on how much and how fast the trader is making money buying the JNJ bonds and hedging them. If the trader is making five basis points on the position and holding it for a day, then a standard deviation of two basis points per day likely represents a reasonable risk–return trade-off. If, on the other hand, the trader is making 1.5 basis points and holding the position for a week, a standard deviation of two basis points per day likely ruins the trade from a risk–return perspective.

This section concludes with an *out-of-sample* analysis of the estimated regression model. Figure 6.2 shows the same regression line as estimated in Table 6.2 and graphed in Figure 6.1. The plus signs, however, are the changes in yields over the period May 17, 2021, to July 19, 2021. The regression model, estimated over the earlier period, January 29, 2021, to May 15, 2021, holds up quite well. In fact, the standard error of the residuals of the out-of-sample data against the original regression line is 1.5 basis points, which is actually smaller than the in-sample equivalent. The trader hedging as of May 14, 2021, cannot, of course, run this analysis. But other

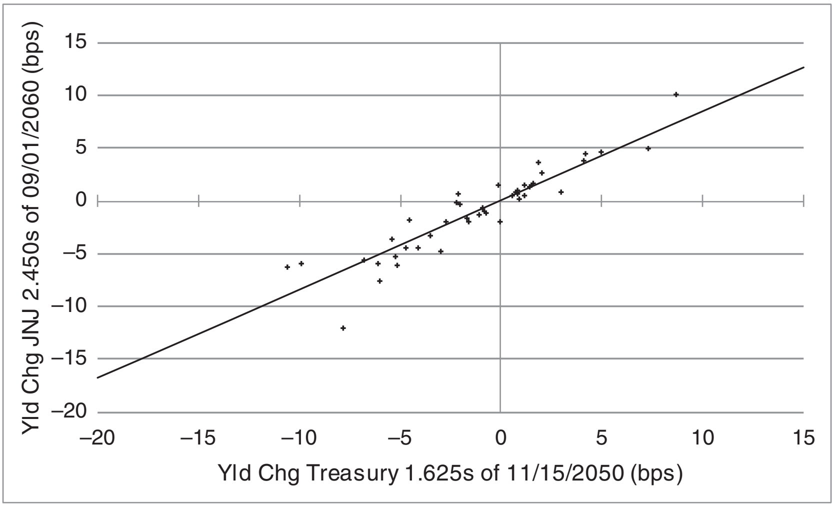

FIGURE 6.2 Yield Changes of the JNJ 2.450s of 09/01/2060 and the Treasury 1.625s of 11/15/2050 from May 17, 2021, to July 19, 2021, and the Regression Line Estimated over the Period January 19, 2021, to May 14, 2021.

out-of-sample tests can be informative. A trader might see how a regression model performed over a period before the estimation period, perhaps a period right before that window or perhaps an even earlier period that might be more likely to resemble the future. In any case, poor out-of-sample performance should raise questions about the stability of the regression coefficients over time and, therefore, the reliability of the resulting hedge.

6.2 TWO-VARIABLE REGRESSION HEDGING

The hedging approaches of Chapter 5 generalize those of Chapter 4 to account for the fact that rates across the term structure are not perfectly correlated. Similarly, two-variable regression hedges account for the fact that changes in a bond's yield might be better explained by changes in the yields of two other bonds, rather than just one, as in a one-variable regression.

To illustrate two-variable regression hedging, consider a relative value trader who believes that yields in the 20-year US Treasury bond sector are too high – or prices too low – relative to the 10- and 30-year sectors. Implementing this trade idea by buying a 20-year bond outright is too risky: if rates increase across the board, the trade loses money even if the trader is right, that is, even if the 20-year bond does outperform 10- and 30-year bonds. But buying a 20-year bond and hedging the interest rate risk by selling a 10-year bond also bears too much risk that is unrelated to the trade idea: if the curve steepens (e.g., 30-year yields increase more than 20-year yields, which increase more than 10-year yields), the trade may lose money even if the 20-year bonds outperform. And, finally, buying a 20-year bond and hedging with a 30-year bond can lose money if the curve flattens even if the 20-year bonds outperform. In practice then, this trade idea is typically implemented with a *butterfly*: buy 20-year bonds and sell both 10- and 30-year bonds: both shorts defend against general rate increases; the 10-year short defends against flattening; and the 30-year short defends against steepening. The trader's problem then becomes to choose the face amount of the 10- and 30-year bonds to sell against, say, \$100 million face amount of the 20-year bond.

In this illustration, the trader chooses the three Treasury bonds listed in Table 6.1: the 1.375s of 11/15/2040 as the 20-year; the 0.875s of 11/15/2030 as the 10-year; and the 1.625s of 11/15/2050 as the 30-year. The two-variable regression model of changes in yields of these bonds is,

$$\Delta y_t^{20} = \alpha + \beta^{10}\Delta y_t^{10} + \beta^{30}\Delta y_t^{30} + \epsilon_t \tag{6.13}$$

where the notation is analogous to that of the one-variable regression. Here there are two slope coefficients, describing how changes in the 20-year yield are related to changes in each of the 10-year and 30-year yields.

Continuing as in the case of one-variable regression, least-squares estimation finds the regression coefficients so as to minimize the sum of the squared residuals,

$$\sum_t \left(\Delta y_t^{20} - \hat{\alpha} + \hat{\beta}^{10}\Delta y_t^{10} + \hat{\beta}^{30}\Delta y_t^{30}\right)^2 \tag{6.14}$$

And, with these estimated coefficients, the predicted change of the 20-year rate is,

$$\Delta \hat{y}_t^{20} = \hat{\alpha} + \hat{\beta}^{10}\Delta y_t^{10} + \hat{\beta}^{30}\Delta y_t^{30} \tag{6.15}$$

Table 6.3 gives the results of the regression, estimated with data from January 29, 2021, to May 14, 2021. The R-squared is quite high relative to that of the single-variable regression, in Table 6.2, in part because two explanatory variables are used, rather than one, and in part because all of the bonds in this regression are Treasuries, whereas the single-variable regression mixes a corporate bond with a Treasury bond. The standard error is also significantly lower here, at 1.15 basis points per day. Again, however, as usual for regressions of this sort, the estimate of $\hat{\alpha}$ is small and not significantly different from zero.

The slope coefficients say that a one-basis-point increase in the 10-year yield increases the 20-year yield by 0.465 basis points, while a one-basis-point increase in the 30-year yield increases the 20-year yield by 0.669 basis points. With 95% confidence intervals for these coefficients of (0.329, 0.601) and (0.535, 0.803), respectively, both coefficients are significantly different from zero; that is, changes in the yields of both bonds are useful in explaining changes in the yield of the 20-year bond.

TABLE 6.3 Regression of Daily Changes in Yields of the Treasury 1.375s of 11/15/2040 on Daily Changes in Yields of the Treasury 0.875s of 11/15/2030 and 1.625s of 11/15/2050, from January 29, 2021, to May 14, 2021.

No. of Observations	**82**	
R-Squared	**94.7%**	
Standard Error	**1.15**	
Regression Coefficients	**Value**	**Std. Error**
Constant ($\hat{\alpha}$)	0.019	0.129
Chg 10yr Treasury Yield ($\hat{\beta}^{30}$)	0.465	0.068
Chg 30yr Treasury Yield ($\hat{\beta}^{10}$)	0.669	0.067

To derive the hedge based on these regression results, write the P&L of the hedged position as,

$$P\&L = -F^{20}\frac{DV01^{20}}{100}\Delta y_t^{20} - F^{10}\frac{DV01^{10}}{100}\Delta y_t^{10} - F^{30}\frac{DV01^{30}}{100}\Delta y_t^{30} \quad (6.16)$$

and then substitute for Δy_t^{20} from (6.15) to see that,

$$\begin{aligned} P\&L = &\left[-F^{20}\frac{DV01^{20}}{100}\hat{\beta}^{10} - F^{10}\frac{DV01^{10}}{100}\right]\Delta y_t^{10} \\ &+ \left[-F^{20}\frac{DV01^{20}}{100}\hat{\beta}^{30} - F^{30}\frac{DV01^{30}}{100}\right]\Delta y_t^{30} \end{aligned} \quad (6.17)$$

Next, to ensure that P&L is zero, under the assumption that the change in the 20-year rate follows the regression model, set each of the terms in brackets in Equation (6.17) equal to zero. Solving,

$$F^{10} = -F^{20}\frac{DV01^{20}}{DV01^{10}}\hat{\beta}^{10} \quad (6.18)$$

$$F^{30} = -F^{20}\frac{DV01^{20}}{DV01^{30}}\hat{\beta}^{30} \quad (6.19)$$

or, in terms of risk weights,

$$\frac{-F^{10} \times DV01^{10}}{F^{20}DV01^{20}} = \hat{\beta}^{10} \quad (6.20)$$

$$\frac{-F^{30} \times DV01^{30}}{F^{20}DV01^{20}} = \hat{\beta}^{30} \quad (6.21)$$

Assuming a trade size of $100 million face amount in the 20-year Treasury, substituting the DV01s of the bonds from Table 6.1 and the results of the regression from Table 6.3, the hedging face amounts and risk weights are $79.44 million and 46.5% for the 10-year, along with $50.69 million and 66.9% for the 30-year. Note that the sum of the risk weights is 113.4%, which means that the DV01 of the hedging position is 13.4% greater than the DV01 of the position being hedged. This follows immediately from the slope coefficients of the regression: a simultaneous one-basis-point change in both the 10- and 30-year yields is associated with a 1.134-basis-point increase in the 20-year yield. Hence, the hedging portfolio requires an extra 13.4% in DV01.

Figure 6.3 compares in-sample and out-of-sample residuals from the regression in Table 6.3. The out-of-sample residuals are very well behaved, in fact, better behaved than the in-sample residuals: the standard error is 1.15 in-sample, and only 0.70 out-of-sample. Traders are not always so fortunate!

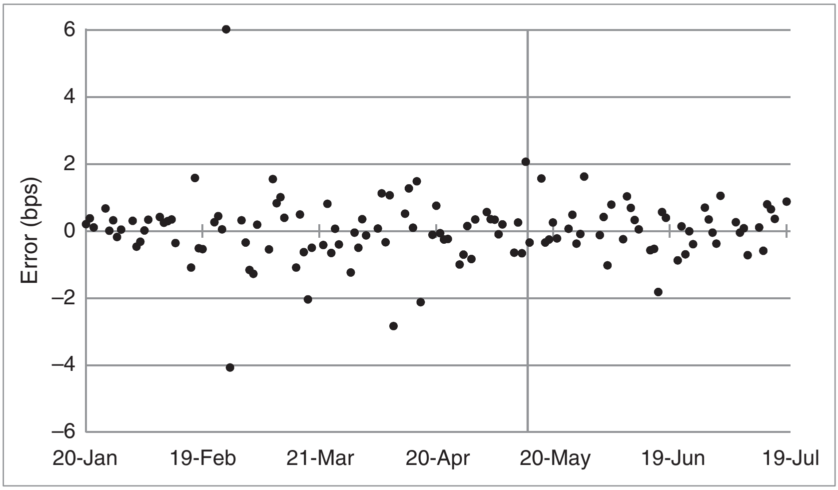

FIGURE 6.3 Residuals Using the Regression Coefficients in Table 6.3, in-Sample – from January 19, 2021, to May 14, 2021 – and Out-of-Sample – from May 17, 2021, to July 19, 2021.

6.3 LEVEL VERSUS CHANGE REGRESSIONS

When estimating regression-based hedges, most practitioners regress changes in yields on changes in yields, as in the previous sections, but some regress yields on yields. Mathematically, in the single-variable case, the level-on-level regression with dependent variable y and independent variable x is,

$$y_t = \alpha + \beta x_t + \epsilon_t \tag{6.22}$$

while the change-on-change regression is,[3]

$$y_t - y_{t-1} = \beta(x_t - x_{t-1}) + \epsilon_t - \epsilon_{t-1} \tag{6.23}$$

$$\Delta y_t = \beta \Delta x_t + \Delta \epsilon_t \tag{6.24}$$

If the assumptions of least-squares estimation, mentioned earlier, hold true for the level model (6.22), then they also hold for the change model (6.24), and least-squares estimates from both specifications are unbiased, consistent, and efficient. If, however, the assumption about the independence of the error terms is violated in either specification, then the least-squares

[3]While it is usual to include a constant term in the change-on-change regression, the constant is omitted here for expositional purposes.

estimates from that specification may not be efficient, but they are still unbiased and consistent.

To discuss the economics behind the assumption that error terms are independent, say that $\alpha = 0$, that $\beta = 1$, and that y and x are the yields of two different bonds. Say further that the yield of the x-bond is constant at 5%, while the yield of the y-bond was 1% yesterday. The level regression, in Equation (6.22), predicts that the yield of the y-bond will be 5% today, despite its having been 1% yesterday. It is more likely, however, that the yield of the y-bond today will be closer to 1% than to 5%, and that the model error today will be closer to its value yesterday, of −4%, than to zero. In other words, the error terms of the level regression are not likely to be independent of each other, but rather persistent, correlated over time, or *serially correlated*.

In this same scenario, because the change in the yield of the x-bond is zero, the change-on-change regression in Equation (6.24) predicts that the change in the yield of the y-bond is zero as well and that its yield remains at 1%. While more plausible than the level-on-level prediction that the yield of the y-bond suddenly jumps to 5%, it is more likely that the yield of the y-bond will gradually trend from its current value of 1% to its model value of 5%. Hence, the error terms in the change-on-change regression are likely to be positive for some time, and, as such, serially correlated.

This discussion suggests an alternate model, which would capture, in the scenario just discussed, that the yield of the y-bond moves gradually from 1% to 5%. In particular, assume the level-on-level model, but with error dynamics,

$$\epsilon_t = \rho\epsilon_{t-1} + \nu_t \tag{6.25}$$

for some $\rho < 1$. In this model, with say, $\rho = 75\%$, yesterday's error of −4% would fall, on average, to an error of 75% times −4%, or −3% today, thus giving an expected y-bond yield today of 2%. In this way, the error structure in Equation (6.25) gradually pushes the yield of the y-bond up from its starting point of 1% to its model value, that is, the 5% yield of the x-bond. The procedure for estimating (6.22) with the error structure in (6.25) is given in many statistical texts.

6.4 REVERSE REGRESSIONS

In Section 6.1, a trader regresses changes in yields of the JNJ 2.450s of 09/01/2060 – with a DV01 of 0.2124 – on changes in yields of the Treasury 1.625s of 11/15/2050 – with a DV01 of 0.1910; obtains a regression coefficient of 0.842; and, against \$100 million of the JNJ bonds, calculates a regression hedge to sell \$100 million × (0.2124/0.1910) × 0.842, or \$93.6 million Treasury bonds.

What if another trader runs the *reverse* regression, that is, regresses changes in yield of the Treasury bond on changes in yield of the JNJ bond? Table 6.4 compares the slope coefficients and standard errors of the original regression and the reverse regression. With a reverse regression $\hat{\beta}$ of 0.921, this second trader hedges \$93.6 million Treasury bonds with \$93.6 million $\times$ (0.1910/0.2124) $\times$ 0.921, or \$82.8 million JNJ bonds.

These hedges are clearly different. The same \$93.6 million of Treasuries are hedged with a different amount of JNJ bonds. Or, in terms of risk weights, both quoted as the DV01 of the Treasury bond position as a percent of the DV01 of the JNJ position, the risk weight of the regression is 84.2%, while the risk weight of the reverse regression is 1/0.921, or 108.6%. Is one of these hedges right and the other wrong?

This question actually reveals the importance of the trader's decision in Section 6.1 to hedge \$100 million face amount of JNJ bonds. Choosing this face amount actually sets the risk of the trade. As shown earlier, the volatility of the hedged position is \$100 million $\times$ 0.2124/100 $\times$ the two-basis-point standard error of the regression, or about \$425,000. However, the risk of the reverse regression, which sets the face amount of the Treasury bonds at \$93.6 million, is \$93.6 million $\times$ 0.1910/100 $\times$ the 2.09 standard error of the reverse regression, or about \$374,000. These trades, therefore, are not comparable.

Choosing to hedge \$100 million face amount of the JNJ bonds, however, is not just about the risk of the hedged position. There are many combinations of positions in the JNJ and Treasury bonds that have the same volatility.[4] For example, a scaled-up reverse regression hedge, with \$106.37 million Treasury bonds and \$88 million JNJ bonds (i.e., \$106.37 million $\times$ (0.1910/0.2124) $\times$ 0.921), has the same \$425,000 volatility as the regression hedge (\$106.37 million $\times$ 0.1910/100 $\times$ 2.09). But while this and other

TABLE 6.4 Regression: Daily Changes in Yields of the JNJ 2.450s of 09/01/2060 on Daily Changes in Yields of the Treasury 1.625s of 11/15/2050. Reverse Regression: Daily Changes in Yields of the Treasury 1.625s of 11/15/2050 on Daily Changes in Yields of the JNJ 2.450s of 09/01/2060. Observations Are from January 29, 2021, to May 14, 2021.

	Regression	Reverse Regression
$\hat{\beta}$	0.842	0.921
Standard Error	2.00	2.09

[4]See Equation (A6.15) in Appendix A6.1, which expresses the variance of the P&L as a quadratic in the DV01s of each position.

positions might have the same volatility, because they do not hold $100 million in JNJ bonds, they are different trades. Most obviously, they do not satisfy the objective of buying $100 million of JNJ bonds from a client. Less obviously, to the extent that the return profile of the JNJ and Treasury bonds differ, different portfolios of the two bonds have different return characteristics as well.

In short, the regression hedge in Section 6.1 minimizes the variance of hedging $100 million face amount of the JNJ bonds. Trades with other objectives, like holding a fixed amount of Treasuries or holding a fixed amount of volatility risk with particular return characteristics, are constructed differently.

6.5 PRINCIPAL COMPONENT ANALYSIS

As mentioned in the introduction to this chapter, regression hedging tends to be *ad hoc*, because the relevant bonds and estimation periods are chosen separately for each application. *Principal component analysis* is useful, by contrast, in providing a single, empirical description of the behavior of the term structure that can be applied across a portfolio of fixed income instruments.

To illustrate PCA, this section uses daily data on fixed versus three-month US Dollar (USD) LIBOR swap rates from June 1, 2020, to July 16, 2021.[5] The data set consists of 13 times series, one for each of the terms from one to 10 years, as well as three with terms of 15, 20, and 30 years. These data can be summarized by the variances or standard deviations of changes in each rate and with their pairwise covariances or correlations. Another way to describe the data, however, is with 13 interest rate factors or components, where each factor represents a particular pattern of changes across the 13 rates. One factor, for example, might represent a simultaneous change of 0.2 basis points in the one-year rate, 0.6 basis points in the two-year rate, 1.2 basis points in the three-year rate, etc., up to 3.7 basis points in the 20-year rate, and 3.8 basis points in the 30-year rate. PCA is a way to construct 13 factors, or *principal components* (PCs), such that they have the following properties:

1. The sum of the variances of the PCs equals the sum of the variances of the individual rates. In this sense, the PCs capture the volatility of the set of interest rates.

[5] LIBOR swaps are being phased out at the time of this writing, but a sufficiently long time series of liquid SOFR swap rates is not yet available for the analysis of this section.

2. The PCs are uncorrelated with each other. While changes in rates of one term are highly correlated with changes in rates of another term, the PCs are constructed so that they are uncorrelated.
3. Subject to (1) and (2), each PC is chosen to have the maximum possible variance given all earlier PCs. Therefore, the first PC explains the largest fraction of the sum of the variances of the rates; the second PC explains the next largest fraction; and so forth.

PCs of rates are particularly useful because of an empirical regularity: the sum of the variances of the first three PCs is usually an overwhelming fraction of the sum of the variances across all rates. Therefore, the variances and covariances of all rates are not necessary to describe how the term structure fluctuates: the structure and volatilities of only three PCs suffice. In the simplified context of three rates, Appendix A6.2 describes the construction of PCs in more detail. The text continues with a discussion of computed PCs for USD LIBOR swaps from the data set described earlier.

Figure 6.4 graphs the first three principal components, while Table 6.5 provides similar information in tabular form. Columns (2) to (4) correspond to the three PC curves in the figure, which can be interpreted as follows. A one standard deviation increase in the "level" PC, given both in Column (2) and the solid line in the figure, is a simultaneous increase in the one-year rate of 0.23 basis points; in the two-year rate of 0.59 basis points; in the 10-year rate of 3.44 basis points; and in the 30-year rate of 3.77 basis points. On a particular day, the change in the term structure might be best explained as a 1.5 standard deviation move in the first PC,

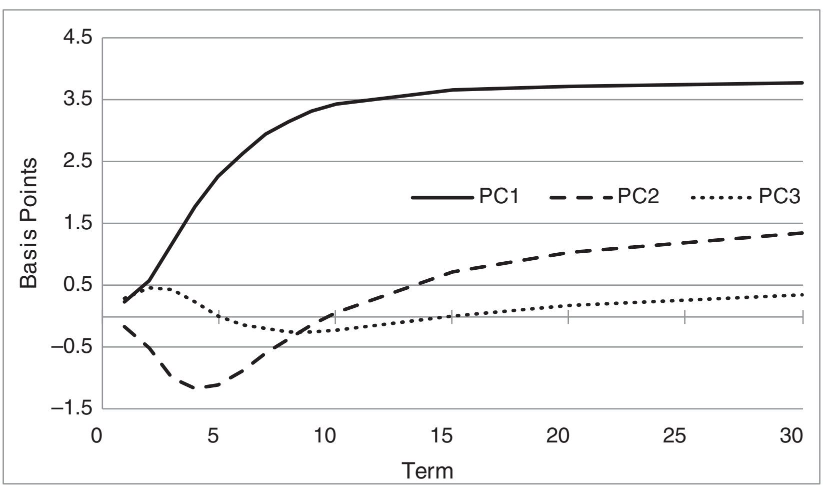

FIGURE 6.4 The First Three Principal Components of USD LIBOR Swap Rates, Estimated from June 1, 2020, to July 16, 2021.

TABLE 6.5 Principal Component Analysis of USD LIBOR Swap Rates from June 1, 2020, to July 16, 2021. Columns (2)-(6) Are in Basis Points; Columns (7)-(10) Are in Percent.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	PCs					% of PC Variance			
Term	Level	Slope	Short Rate	PC Vol	Total Vol	Level	Slope	Short Rate	(5)/(6)
1	0.23	−0.16	0.29	0.41	0.55	32.7	15.0	52.3	74.54
2	0.59	−0.51	0.47	0.91	0.93	42.3	31.5	26.2	97.47
3	1.18	−0.99	0.42	1.59	1.60	54.5	38.5	7.0	99.63
4	1.77	−1.16	0.23	2.13	2.13	69.1	29.7	1.2	99.72
5	2.28	−1.12	0.02	2.54	2.54	80.6	19.4	0.0	99.78
6	2.64	−0.89	−0.13	2.79	2.79	89.7	10.1	0.2	99.91
7	2.94	−0.60	−0.20	3.01	3.01	95.6	4.0	0.4	99.97
8	3.14	−0.36	−0.24	3.17	3.17	98.2	1.3	0.6	99.96
9	3.31	−0.13	−0.25	3.32	3.32	99.3	0.2	0.6	99.92
10	3.44	0.07	−0.23	3.44	3.45	99.5	0.0	0.4	99.87
15	3.65	0.71	0.00	3.72	3.72	96.3	3.7	0.0	99.98
20	3.73	1.02	0.17	3.87	3.87	92.8	7.0	0.2	99.98
30	3.77	1.35	0.36	4.02	4.02	87.9	11.3	0.8	99.93
Total	9.99	2.92	0.96	10.45	10.47	91.3	7.8	0.8	99.84

that is, as adding 1.5 times each element of the first PC to corresponding swap rates. On another day, the change in the term structure might best be described as a −0.75 standard deviation move in the first component, that is, as subtracting 0.75 times each element of the first PC from current rates. In any case, the first PC has been traditionally called the "level" component because it has typically represented an approximately parallel shift over much of its range. In the empirical results presented here, however, the component is not particularly level for terms from one to seven years.

A one standard deviation increase in the "slope" PC, given both in Column (3) of the table and the dashed line in the figure, is a simultaneous fall in the one-year rate of 0.16 basis points; a fall in the two-year rate of 0.51 basis points; an increase in the 10-year rate of 0.07 basis points; and an increase in the 30-year rate of 1.35 basis points. This PC is said to represent a "slope" change in rates because shorter-term rates fall while longer-term rates increase, or *vice versa*.

Lastly, a one standard deviation increase in the "short-rate" PC, given both in Column (4) of the table and the dotted line in the figure, is a simultaneous small increase of very short-term rates; a small decrease in intermediate-term rates, and a small increase in long-term rates. Because of its shape across terms, this PC is often named "curvature" as well, but, in light of the full discussion in this section, this PC is particularly useful for adding explanatory power to variations in shorter-term rates.

To reiterate the sense in which the PCs describe changes in the term structure, on a given day, changes across terms might be described – picking numbers at random – as the combination of: a +1.5 standard deviation change in the first PC; a −0.4 standard deviation change in the second PC; and a −1.8 standard deviation change in the third PC. The term structure at the end of that day, therefore, would approximately equal the term structure at the end of the previous day plus the contributions from the multiples of each of the three PCs. In this way, as explained shortly, these three PCs can indeed explain an overwhelmingly large proportion of realized term structure volatility.

The small values of the PCs at very short-term rates reflect the low volatility of these rates. In the current financial environment, with the Federal Reserve promising to keep short-term rates low for an extended period of time, current economic shocks are not envisioned as impacting short-term rates until some time in the future. As a result, economic volatility is not reflected in very short-term rates but seeps gradually into intermediate- and longer-term rates as expectations of reactions to future Federal Reserve policy actions.[6]

Column (5) of Table 6.5 gives the combined standard deviation or volatility from the three principal components for each rate, while Column (6) gives the total, empirical volatility of each rate over the sample period. For the five-year rate, for example, recalling that PCs are, by construction, uncorrelated, the volatility from the three PCs is,

$$\sqrt{2.28^2 + (-1.12)^2 + 0.02^2} = 2.54 \tag{6.26}$$

The total volatility of the five-year rate in the sample is also, to two decimal places, 2.54, but Column (10) – using more decimal places than shown in Columns (5) and (6) – reports that the ratio of five-year PC volatility to total volatility is 99.78%. Hence, the empirical variation of the five-year rate is almost completely explained by the first three PCs. Considering Column (10) as a whole, three PCs explain over 99% of the variation of rates of all

[6]For many years before the financial crisis of 2007–2009, the first PC was hump-shaped, increasing to a peak at about five years or so, and then declining gradually over longer terms. This shape was interpreted as the Federal Reserve pegging very short-term rates but responding in the relatively near term to economic volatility. Also, because current economic volatility affects views on longer-term rates less and less with term, volatility eventually begins to decline with term. Current Federal Reserve policy, however, as discussed in the text, seems to have changed dramatically this empirical feature of rates markets.

terms greater than three years, 97.47% of the variation in the two-year rate, and 74.54% of the variation in the one-year rate. Hence, although there are 13 rates in the data series, three factors alone – three fixed combinations of changes in rates across terms – go a very long way in explaining the variation in all 13 rates. This is possible, intuitively, because changes in rates across terms are highly correlated; that is, nowhere near 13 factors are actually necessary to explain the variation in 13 rates. The performance of the three factors is less impressive, however, for rates of the shortest term.

Columns (7) through (9) of Table 6.5 give the variance explained by each of the first three PCs as a fraction of the total variance explained by those three PCs. For the two-year rate, for example, those fractions are calculated as follows,

$$\frac{0.5916^2}{0.9091^2} = 42.3\% \tag{6.27}$$

$$\frac{(-0.5101)^2}{0.9091^2} = 31.5\% \tag{6.28}$$

$$\frac{0.4650^2}{0.9091^2} = 26.2\% \tag{6.29}$$

Note that, to avoid confusion, the values in these equations are reported to greater accuracy than those in the table.

Looking at Columns (7) to (9) across terms, the level PC dominates the other two as the main contributor to variations in the term structure. The short-rate PC, and then the slope PC, are significant contributors at the short end, however, as is the slope PC for the longest rates. These finding have significant implications for risk management. A trader of eight- to 10-year swaps, or perhaps even seven- to 15-year swaps, can defend using the one-factor metrics and hedging approaches of Chapter 4 and one-variable regression hedging described in this chapter: according to Table 6.5, in this range of terms, term structure variation can be well described by a one-factor model, like the level PC. On the other hand, traders in the three- to six-year sector or the long end might very well require two factors, while traders in the very short end may not be comfortable without a three-factor framework.

The last row of Table 6.5 computes the various statistics just discussed across the whole term structure of rates. More specifically, Columns (2) to (6) give the square root of the sum of variances across terms, and Columns (7) to (9) give the respective ratios for these totals. While the sum of variances is not a particularly interesting economic quantity – it does not represent the variance of any particularly interesting portfolio – the last row of the table does summarize two overall results of the PCA. First, 99.84% of the volatility of the 13 rates in the study is explained by the first three PCs. Second, the level PC explains over 90% of that variance, the slope PC about 8%, and the short-rate PC less than 1%.

The overall lessons from PCA are often similar across global markets. Estimated over the same time period as the USD PCs in Figure 6.4, Figures 6.5 and 6.6 graph the first three PCs of British Pound Sterling (GBP) LIBOR and Euribor swap rates, respectively.[7] The GBP PCs are extremely similar to their USD equivalents, with respect to both shape and magnitude. The biggest difference seems to be that the slope PC in USD is more volatile. The shapes of the EUR PCs are qualitatively similar to those in USD and GBP, but volatility in EUR is significantly lower. The EUR level PC, for example, flattens in the long end at a bit above 2.5 basis points per day, whereas the USD and GBP level PCs flatten at over 3.5. This lower volatility might be explained by the current aggressiveness of the European Central Bank, relative to the Federal Reserve and the Bank of England, to keep short-term rates low over an extended period of time.

Hedges based on PCA are constructed like the multi-factor approaches of Chapter 5. Using the current term structure, calculate the current price of the portfolio being hedged; shift the current term structure by each PC, in turn, to get new term structures and new portfolio prices; with these new prices, calculate a portfolio ’01 with respect to each PC; and find a portfolio of hedging securities that neutralizes these ’01s.

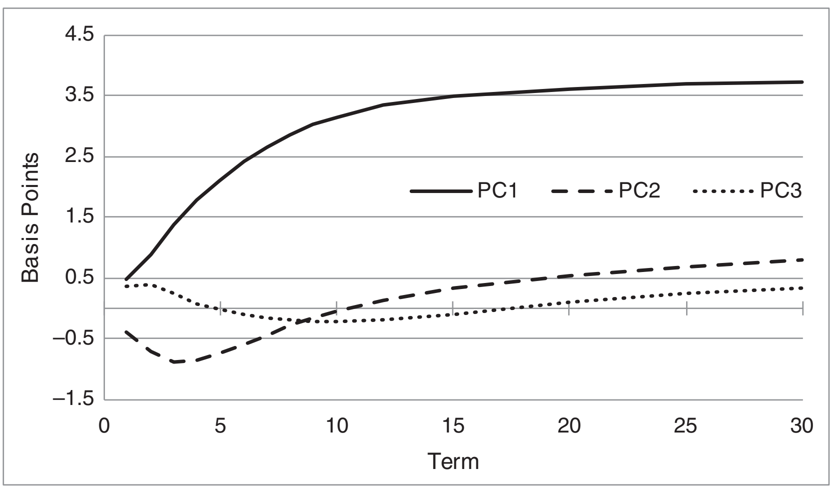

FIGURE 6.5 The First Three Principal Components of GBP LIBOR Swap Rates, Estimated from June 1, 2020, to July 16, 2021.

[7]For a description of these floating-rate indexes, see Chapter 12.

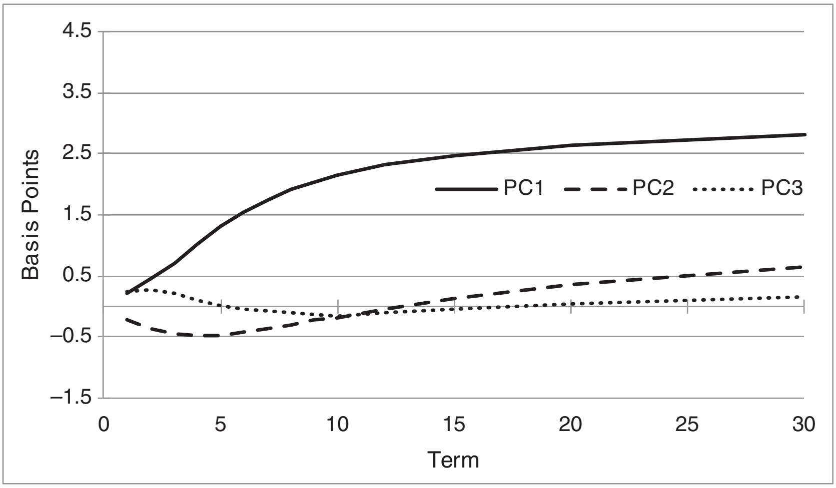

FIGURE 6.6 The First Three Principal Components of Euribor Swap Rates, Estimated from June 1, 2020, to July 16, 2021.

TABLE 6.6 USD LIBOR Par Swap Rates and DV01s, as of July 16, 2021, and PC Elements from Table 6.5. Rates Are in Percent, and PC Elements Are in Basis Points.

Term	Rate	DV01	Level	Slope	Short Rate
10	1.303	0.09347	3.44	0.07	−0.23
15	1.501	0.13387	3.65	0.71	0.00
20	1.596	0.17064	3.73	1.02	0.17
30	1.646	0.23600	3.77	1.35	0.36

This section illustrates hedging with PCA in a context described earlier, namely, hedging a relative value 10s-20s-30s butterfly, but this time in swaps. Specifically, a trader believes that USD 20-year swap rates are too high relative to 10- and 30-year swap rates, and plans, therefore, to receive fixed in 20s and pay in 10s and 30s.[8] Par swap rates and DV01s of par swaps are given in Table 6.6. Also, corresponding to rates of the listed terms, the table gives the elements of each of the three PCs from Table 6.5. By definition, a one standard deviation shift in each PC changes these swap rates by these PC elements. The inclusion of the 15-year swap rate is discussed later.

Assume that the trader plans to receive fixed on 100 notional amount of the 20-year swap and on F^{10} and F^{30} notional amount of 10- and 30-year

[8]From the material presented in Chapter 2, the reader can think of this trade as "buying" 20-year bonds and "selling" 10- and 30-year bonds.

swaps, respectively. Paying fixed is reflected, in this notation, with negative notional amounts. In any case, from the data in Table 6.6, the exposure of this overall relative value portfolio is hedged against a one standard deviation shift of the level and slope PCs, respectively, if the following equations obtain,

$$-F^{10}\frac{0.09347}{100}\times 3.44 - F^{30}\frac{0.23600}{100}\times 3.77 - 100\frac{0.17064}{100}\times 3.73 = 0 \tag{6.30}$$

$$-F^{10}\frac{0.09347}{100}\times 0.07 - F^{30}\frac{0.23600}{100}\times 1.35 - 100\frac{0.17064}{100}\times 1.02 = 0 \tag{6.31}$$

The first two terms of Equation (6.30) give the change in the value of the hedge position under a one standard deviation shift of the first PC, that is, a shift of +3.44 basis points in the 10-year and +3.77 basis points in the 30-year swap rate. The third term gives the change in the value of the position being hedged under the same PC shift, which is +3.73 basis points in the 20-year swap rate. Note that the negative signs indicate that receiving fixed (i.e., positive notional amounts) when rates increase results in position losses. The equation as a whole, therefore, sets the total position gain or loss under a one standard deviation shift of the first PC equal to zero. Equation (6.31) can be interpreted similarly, but under a one standard deviation shift of the second PC. Note, of course, that if these two equations hold for a one standard deviation shift, they hold for any size shift: to see this, simply multiply both sides of each equation by the intended number of standard deviations.

Solving Equations (6.30) and (6.31) reveals that $F^{10} = -49.56$ and $F^{30} = -53.60$. Or, in terms of risk weights relative to the DV01 of the 20-year swap,

$$\frac{49.56\times\frac{0.09347}{100}}{0.17064} = 27.1\% \tag{6.32}$$

$$\frac{53.60\times\frac{0.23600}{100}}{0.17064} = 74.1\% \tag{6.33}$$

Intuitively, most of the risk of the 20-year swap – 74% – is hedged with 30-year swaps, because the exposures of 20-year swaps to both the level and slope PCs more closely resemble those of 30-year swaps than of 10-year swaps. Note also that the sum of the risk weights is 101.2%, so that the DV01 of the hedge position is greater than the DV01 of the position being hedged. Only under the assumption of parallel shifts do the risk weights always sum to 100%. In the present case, more DV01 risk is needed in the hedge because the hedge includes a significant amount of 10-year swaps, which are much less sensitive to the level and slope shifts than the 20-year swaps.

In this illustration, the trader chooses to hedge with 10- and 30-year swaps. But, with only two hedging securities, the risks of only two PCs can be hedged. What is the risk of the hedged position just derived to the next most important PC, that is, the short-rate or curvature PC? Following the same logic as in Equations (6.30) and (6.31), the exposure of the hedged position to the third PC (adding a significant digit to avoid confusion) is,

$$\begin{aligned} &-(-49.6)\frac{0.09347}{100}\times(-0.228)-(-53.6)\frac{0.23600}{100}\times 0.360 \\ &-100\frac{0.17064}{100}\times(0.166)=0.007 \end{aligned} \tag{6.34}$$

which is less than one cent per 100 face amount. The trader might very well decide, therefore, that it is not worth the transaction costs of trading an additional swap to hedge out this residual risk from the third PC. Also, because this is a relative value trade, the trader wants to pay fixed only in swaps with rates that are believed to be too low. In any case, if hedging out the residual risk is desired, a 15-year swap can be added to the hedging portfolio; an equation for exposure to the third PC can be added to Equations (6.30) and (6.31); and, using the data from Table 6.6, the notional amounts for the 10-, 15-, and 30-year swaps can be determined. This is left as an exercise for the reader.

CHAPTER 7

Arbitrage Pricing with Term Structure Models

Principal components analysis reveals that the term structure of interest rates is determined by relatively few factors or random processes. Therefore, assumptions about how these few factors evolve over time, combined with arbitrage arguments, can deliver strong predictions about the prices and interest rate sensitivities of bonds and other *interest rate contingent claims* (i.e., securities with cash flows that depend on interest rates, like bond options). Formulating assumptions about the evolution of interest rate factors, pricing fixed income securities, and determining hedge ratios comprise the art and science of *term structure models*.

Term structure models are presented in three chapters. This chapter uses a very simple setting to show how assumptions about the evolution of the short-term rate over time allows for the arbitrage pricing of bonds of all maturities and of interest rate contingent claims. *Option-adjusted spread (OAS)* is introduced both as a metric of a security's mispricing relative to a model and as the spread that can be earned – if the model is correct – by trading that security. Chapter 8 shows how the shape of the term structure is determined by: expectations about future short-term rates, the risk premium required by investors to bear interest rate risk, and convexity, whose effect is a result of interest rate volatility. Chapter 9 then illustrates the art of modeling the evolution of short-term rates by presenting two term structure models: the classic Vasicek model and the two-factor Gauss+ model, which has proven popular in industry for both relative value and macro-style trading.

7.1 RATE AND PRICE TREES

Assume that the six-month and one-year spot rates are 2% and 2.15%, respectively. Taking these market rates as given is equivalent to taking the prices of a six-month bond and a one-year-bond as given. Securities with assumed prices are called *underlying securities* to distinguish them from the contingent claims priced by arbitrage arguments.

Next, assume that six months from now the six-month rate is either 2.50% or 1.50% with equal probability. This very strong assumption is depicted at the top of Figure 7.1 by means of a *binomial tree*, where "binomial" means that only two future values are possible. The columns in the tree represent dates. The six-month rate is 2% today, which is called date 0. Six months from now, on date 1, there are two possible outcomes or *states of the world*. The 2.50% state is called the *upstate* while the 1.50% state is called the *downstate*.

Given the current term structure of spot rates (i.e., the current six-month and one-year rates), trees can be computed for the prices of six-month and one-year zero coupon bonds. The price tree for \$1,000 face value of the six-month zero, depicted at the bottom left of Figure 7.1, shows that the date-0 price is \$1,000/(1 + 0.02/2) = \$990.099. (For readability, currency symbols are not included in price trees.) Note that, in a tree for the value of a particular security, the maturity of the security falls over time. On date 0 of the tree just discussed, the security is a six-month zero, while on date 1 the security is a maturing zero.

The price tree for \$1,000 face value of a one-year zero is depicted at the bottom right of Figure 7.1. The three prices on date 2 are all \$1,000, which is the face value of the one-year zero. The two prices on date 1 are found by discounting this certain \$1,000 at the then-prevailing six-month rate. Hence, the date 1 upstate price is \$1,000/(1 + 0.025/2), or \$987.654, and

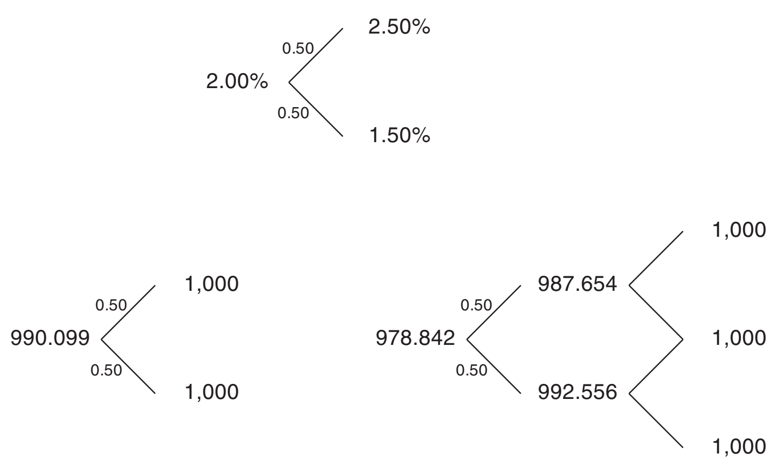

FIGURE 7.1 Pricing Six-Month and One-Year Zero Coupon Bonds with a Binomial Rate Tree.

the date 1 downstate price is \$1,000/ (1 + 0.015/2), or \$992.556. Finally, the date 0 price is computed using the given, date 0, one-year rate of 2.15%: $\$1,000/(1 + 0.0215/2)^2$, or 978.842.

The probabilities of moving up or down the tree may be used to compute average or expected values. As of date 0, the expected value of the one-year zero price on date 1 is,

$$0.5 \times \$987.654 + 0.5 \times \$992.556 = \$990.105 \tag{7.1}$$

Discounting this expected value to date 0, at the date 0, six-month rate gives an *expected discounted value* of,

$$\frac{0.5 \times \$987.654 + 0.5 \times \$992.556}{1 + \frac{.02}{2}} = \$980.302 \tag{7.2}$$

Note that the one-year zero's expected discounted value of \$980.302 is not equal to its market price of \$978.842. These two numbers need not be equal, because investors do not price securities by expected discounted value. Over the next six months, the one-year zero is a risky security, worth \$987.654 half of the time and \$992.556 the other half of the time, for an average or expected value of \$990.105. If investors do not like this price uncertainty, they would prefer a security worth \$990.105 on date 1 with certainty. More specifically, a security worth \$990.105 with certainty after six months would sell for \$990.105/(1 + .02/2), or \$980.302, as of date 0. By contrast, investors penalize the risky one-year zero coupon bond with an average price of \$990.105 in six months by pricing it today at \$978.842. Chapters 8 and 9 elaborate further on investor *risk aversion*.

7.2 ARBITRAGE PRICING OF DERIVATIVES

This section prices an interest rate contingent claim or derivative, in particular, a call option that expires in six months to purchase \$1,000 face value of a then six-month zero at \$990. Figure 7.2 starts the price tree for this call option based on the rates and prices in Figure 7.1. If on date 1 the six-month rate is 2.50%, and a six-month zero sells for \$987.654, the right to buy that zero at \$990 is worthless. On the other hand, if the six-month rate is 1.50%, and the price of a six-month zero is \$992.556, then the right to buy the zero at \$990 is worth \$992.556 – \$990, or \$2.556. This description of the option's terminal payoffs emphasizes the contingent claim nature of the option: its value depends on interest rates through the value of an underlying bond.

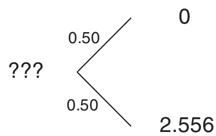

FIGURE 7.2 Pricing a 990 Six-Month Call Option on a Six-Month Zero Coupon Bond.

Chapter 1 showed that a security is priced by arbitrage by finding and pricing its replicating portfolio. In that context, because all bond cash flows are fixed or constant, the construction of the replicating portfolio is relatively simple. The present context is more difficult, because cash flows do depend on the level of rates, and the replicating portfolio must replicate the contingent claim for any possible interest rate scenario.

To price the call option of this section by arbitrage, construct a portfolio on date 0 of underlying securities, namely six-month and one-year zero coupon bonds, such that the portfolio is worth \$0 in the upstate on date 1 and \$2.556 in the downstate. Let $F^{.5}$ and F^{1} be the face values of six-month and one-year zeros in this replicating portfolio, respectively, and recall that the possible values of these bonds on date 1 are shown in Figure 7.1. These face amounts, therefore, must satisfy the following two equations,

$$F^{.5} + .987654F^{1} = \$0 \tag{7.3}$$

$$F^{.5} + .992556F^{1} = \$2.558 \tag{7.4}$$

Equation (7.3) may be interpreted as follows. In the upstate, the value of the replicating portfolio's now maturing six-month zero is its face value. The value of the once one-year zeros, now six-month zeros, is .987654 per dollar face value. Hence, the left-hand side of Equation (7.3) denotes the value of the replicating portfolio in the upstate. This value must equal \$0, the value of the option in the upstate. Similarly, Equation (7.4) sets the value of the replicating portfolio in the downstate equal to the value of the option in the downstate.

Solving Equations (7.3) and (7.4) gives $F^{.5} = -\$515.0000$ and $F^{1} = \$521.4375$. In words, the option can be replicated by buying \$521.4375 face value of one-year zeros and shorting \$515.0000 face amount of six-month zeros on date 0. Therefore, by the law of one price, the price of the option equals the price of the replicating portfolio, which, using the bond prices given earlier, is equal to,

$$\begin{aligned} .990099F^{.5} + .978842F^{1} &= -.990099 \times \$515.0000 + .978842 \times \$521.4375 \\ &= \$0.504 \end{aligned} \tag{7.5}$$

Recall that pricing based on the law of one price is enforced by arbitrage. If the price of the option were less than \$0.504, arbitrageurs could buy the option, short the replicating portfolio, keep the difference, and have no future liabilities. Similarly, if the price of the option were greater than \$0.504, arbitrageurs could short the option, buy the replicating portfolio, keep the difference, and, once again, have no future liabilities. Thus, ruling out profits from riskless arbitrage implies an option price of \$0.504.

It is important to emphasize that the option cannot be priced by expected discounted value, which gives an option price of,

$$\frac{.5 \times \$0 + .5 \times \$2.555831}{1 + \frac{.02}{2}} = \$1.2653 \tag{7.6}$$

The true option price is lower, because investors dislike the risk of the call option and, as a result, will not pay as much as its expected discounted value. Put another way, the risk penalty implicit in the call option price is inherited from the risk penalty of the one-year zero, that is, from the property that the price of the one-year zero is less than its expected discounted value. Once again, the pricing of risk is discussed in the next two chapters. While this section illustrates arbitrage pricing with a call option, it should be clear that the framework can be used to price any security with cash flows that ultimately depend on the six-month rate. For example, because the price of a five-year bond over time depends on the evolution of the six-month rate, an option on that five-year bond can be priced in this framework as well.

A remarkable feature of arbitrage pricing is that the probabilities of up- and down-moves never enter into the calculation of the arbitrage price. See Equations (7.3) through (7.5). The explanation for this somewhat surprising result follows from the principles of arbitrage. Arbitrage pricing requires that the value of the replicating portfolio be the same as the value of the option in both the up- and the down-states. Therefore, the composition of the replicating portfolio is the same whether the probability of the upstate is 20%, 50%, or 80%. But if the composition of the portfolio does not depend directly on the probabilities, and if the prices of the securities in the portfolio are given, then the price of the replicating portfolio and the price of the option cannot depend directly on the probabilities either.

Despite the fact that the option price does not depend directly on the probabilities, these probabilities must have some impact on the option price. After all, as it becomes more and more likely that rates will rise to 2.50% and that bond prices will be low, the value of options to purchase bonds must fall. The resolution of this apparent paradox is that the option price depends indirectly on the probabilities through the price of the one-year zero. Were the probability of an up move to increase suddenly, the current value of a one-year zero would decline. And since the replicating portfolio is long one-year zeros, the value of the option would decline as well. In summary, a derivative like an option depends on the probabilities only through current bond prices. Given bond prices, however, probabilities are not needed to derive prices determined by arbitrage.

7.3 RISK-NEUTRAL PRICING

Risk-neutral pricing is a technique that modifies an assumed interest rate process, like the one assumed at the start of this chapter, so that any contingent claim can be priced without having to construct and price its replicating portfolio. Because the technique requires that the original interest rate process be modified only once, and because this modification requires no more effort than pricing a single contingent claim by arbitrage, risk-neutral pricing is an extremely efficient way to price many contingent claims under the same assumed rate process.

In the example of this chapter, the price of a one-year zero does not equal its expected discounted value: its price is \$978.842, computed from the given one-year spot rate of 2.15%, while its expected discounted value is \$980.302, as derived in Equation (7.2). The probabilities of 0.5 for the up- and down-states are the assumed *true* or *real-world* probabilities. But there are other probabilities, called *risk-neutral* probabilities, which do cause the expected discounted value to equal the market price. To find these probabilities, let the risk-neutral probabilities in the up- and down-states be p and $(1-p)$, respectively. Then, solve the following equation,

$$\frac{\$987.654p + \$992.556(1-p)}{\left(1+\frac{.02}{2}\right)} = \$978.842 \tag{7.7}$$

to find that $p = .8009$. Hence, under the risk-neutral probabilities of .8009 and .1991, the expected discounted value does equal the market price.

Risk-neutral probabilities can also be described in terms of the *drift* in interest rates. Under the true probabilities, there is a 50% chance that the six-month rate rises from 2% to 2.50%, and a 50% chance that it falls from 2% to 1.50%. Hence, the expected change in the six-month rate, or the drift of the six-month rate, is zero. Under the risk-neutral probabilities, there is an 80.09% chance of a 50-basis point increase and a 19.91% chance of a 50-basis point decrease, for an expected change of 30.09 basis points. Hence, the drift of the six-month rate under these probabilities is 30.09 basis points.

As pointed out in the previous section, the expected discounted value of the option payoff is \$1.2653, while the arbitrage price is \$0.504. But if expected discounted value were computed using the risk-neutral probabilities, the resulting option value would equal its arbitrage price,

$$\frac{.8009 \times \$0 + .1991 \times \$2.555831}{\left(1 + \frac{.02}{2}\right)} = \$0.504 \tag{7.8}$$

The fact that the arbitrage price of the option equals its expected discounted value under the risk-neutral probabilities is not a coincidence. In general, to value contingent claims by risk-neutral pricing, proceed as follows. First, find the risk-neutral probabilities that equate the prices of the underlying securities to their expected discounted values. (In the simple example here, the only risky, underlying security is the one-year zero.) Second, price the contingent claim by expected discounted value under these risk-neutral probabilities. The remainder of this section describes intuitively why risk-neutral pricing works. Since the argument is a bit complex, it is broken up into four steps:

1. Given trees for the underlying securities, the price of a security that is priced by arbitrage does not depend on investors' risk preferences. The reasoning is as follows.

   A security is priced by arbitrage if its cash flows can be replicated by some portfolio of underlying securities. Under the assumed process for interest rates in this chapter, the bond option is priced by arbitrage. By contrast, it is unlikely that a specific common stock can be priced by arbitrage, because no portfolio of underlying securities can mimic the idiosyncratic fluctuations of a single common stock's market value.

   If a security is priced by arbitrage, and if everyone agrees on the price evolution of the underlying securities, then everyone agrees on the replicating portfolio. In the option example, both an extremely risk-averse, retired investor and a professional gambler agree that a portfolio of \$521.4375 face of one-year zeros and -\$515.0000 face of six-month zeros replicates the option. And because they agree both on the composition of the replicating portfolio and on the prices of the underlying securities, they must also agree on the price of the option.
2. Imagine an economy that has the same current bond prices and possible future values of the six-month rate as the true economy. The imaginary economy is different, however, in that its investors are risk neutral. Unlike investors in the true economy, then, investors in the imaginary economy do not penalize securities for risk: they price securities by expected discounted value. In particular, under the probabilities in the imaginary economy, the expected discounted value of the one-year zero equals its market price. But, by Equation (7.7), the expected discounted value of the one-year zero does equal its market

price under the risk-neutral probabilities of .8009 and .1991. Hence, these risk-neutral probabilities are the probabilities in the imaginary economy.

3. The price of the option in the imaginary economy, like any other security in that economy, is computed by expected discounted value. Since the probability of the upstate in that economy is .8009, the price of the option in that economy is given by Equation (7.8) and is \$0.504.
4. Step 1 implies that, given the prices of the six-month and one-year zeros, as well as possible values of the six-month rate, the price of an option does not depend on investor risk preferences. Therefore, because the real and imaginary economies have the same bond prices and the same possible values for the six-month rate, the option price must be the same in both economies. In particular, the option price in the real economy must also equal \$0.504. More generally, the price of a derivative in the real economy may be computed by expected discounted value under the risk-neutral probabilities.

7.4 ARBITRAGE PRICING IN A MULTI-PERIOD SETTING

Maintaining the binomial assumption, Figure 7.3 extends the tree from the previous section for another six months. This tree is called a *recombining tree*, because an up-move followed by a down-move, to the up–down state, lands in the same place as a down-move followed by an up-move, to the down–up state. Trees for which this is not the case are said to be *nonrecombining*. While nonrecombining trees might represent economically reasonable dynamics, they tend to be avoided as difficult or even impossible to implement. After six months there are two possible states, after one year there are four, and after N semiannual periods there are 2^N possibilities. A tree with enough semiannual steps to price 10-year securities has, in its rightmost column alone, over 500,000 nodes, and to price 20-year securities, over 500 billion. Furthermore, as discussed later in the chapter, it is often desirable to reduce the time interval between dates substantially. In short, even with modern computers, trees that grow this quickly are computationally unwieldy. This does not mean that the effects that seem to give rise to nonrecombining trees – like volatilities that change across states – cannot be modeled. It does mean, however, that such effects have to be implemented in more efficient ways.

Returning to the recombining format, as trees grow it becomes convenient to develop a notation with which to refer to particular nodes. One convention is as follows. The dates, represented by columns of the tree, are numbered from left to right starting with 0. The states, represented by rows of the tree, are numbered from bottom to top, also starting from 0. For example, in Figure 7.3, the six-month rate on date 2, state 0 is 1%. The six-month rate on state 1 of date 1 is 2.50%.

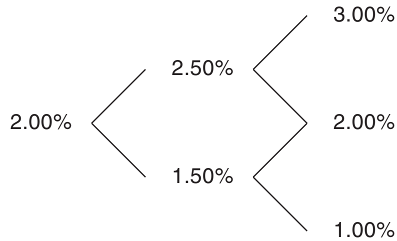

FIGURE 7.3 A Recombining Binomial Rate Tree.

Continuing where the option example left off, having derived the risk-neutral tree for pricing a one-year zero, the goal is to extend the tree to price a 1.5-year zero assuming that the 1.5-year spot rate is 2.25%. Ignoring the probabilities for a moment, several nodes of the 1.5-year zero price tree can be written down immediately, as shown in Figure 7.4. On date 3, the zero with an original term of 1.5 years matures and is worth its face value of \$1,000. On date 2, the value of the then six-month zero equals its face value discounted for six months at the then-prevailing spot rates of 3%, 2%, and 1%, in states 2, 1, and 0, respectively,

$$\frac{\$1{,}000}{1+\frac{.03}{2}} = \$985.22 \tag{7.9}$$

$$\frac{\$1{,}000}{1+\frac{.02}{2}} = \$990.10 \tag{7.10}$$

$$\frac{\$1{,}000}{1+\frac{.01}{2}} = \$995.02 \tag{7.11}$$

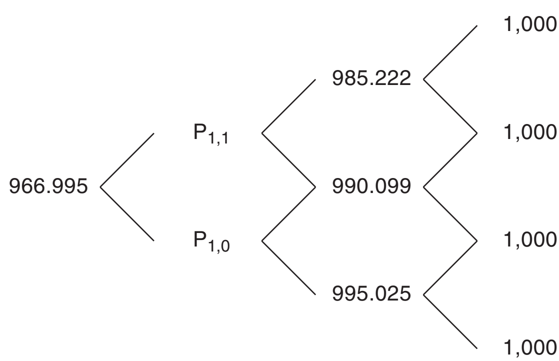

FIGURE 7.4 Price Tree for a 1.5-Year Zero Coupon Bond.

Finally, on date 0, the 1.5-year zero equals its face value discounted at the given, 1.5-year spot rate,

$$\frac{\$1{,}000}{\left(1+\frac{.0225}{2}\right)^3} = \$966.9954 \tag{7.12}$$

The prices of the zero on date 1 in states 1 and 0 are denoted in Figure 7.4 by $P_{1,1}$ and $P_{1,0}$, respectively. These one-year zero prices are not known at this point.

The previous section showed that the risk-neutral probability of an up-move on date 0 is 0.8009. Letting q be the risk-neutral probability of an up-move on date 1, and, for the purposes of this section, making the simplifying assumption that the probability of moving up from state 0 is the same as the probability of moving up from state 1, the resulting tree is shown in Figure 7.5.

By definition, the expected discounted value under risk-neutral probabilities recovers market prices. With respect to the 1.5-year zero price on date 0, this requires that,

$$\frac{.8009P_{1,1} + .1991P_{1,0}}{1+\frac{.02}{2}} = \$966.995 \tag{7.13}$$

And with respect to the prices of a then one-year zero on date 1,

$$P_{1,1} = \frac{\$985.222q + \$990.099(1-q)}{1+\frac{.025}{2}} \tag{7.14}$$

$$P_{1,0} = \frac{\$990.099q + \$995.025(1-q)}{1+\frac{.015}{2}} \tag{7.15}$$

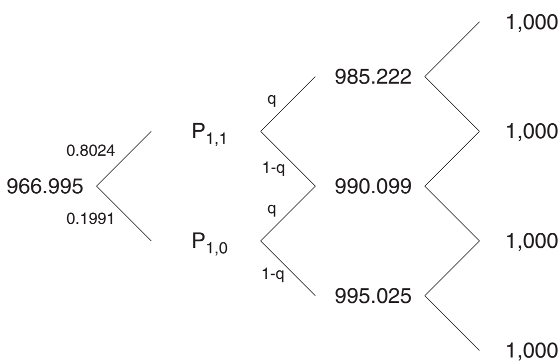

FIGURE 7.5 Price Tree for a 1.5-Year Zero Coupon Bond, with Probabilities.

Substituting Equations (7.14) and (7.15) into Equation (7.13) results in a linear equation in the one unknown, q, which can be solved to find that $q = 0.6520$. Therefore, the risk-neutral interest rate process is summarized by the tree in Figure 7.6. Furthermore, any contingent claim that depends on the six-month rate in six months and in one year may be priced by computing its discounted expected value along this tree. An example is given in the next section.

The difference between the true and risk-neutral probabilities may once again be described in terms of drift. From dates 1 to 2, the drift under the true probabilities is zero. Under the risk-neutral probabilities, the drift is computed from a 65.20% chance of a 50-basis-point increase in the six-month rate and a 34.80% chance of a 50-basis-point decline in the rate. These numbers give a drift or expected change of 15.20 basis points.

Substituting $q = 0.6520$ back into Equations (7.14) and (7.15) completes the tree for the price of the 1.5-year zero, which is shown in Figure 7.7. It follows immediately from this tree that the one-year spot rate in six months may be either 2.5754% or 1.5754%, because,

$$\$974.735 = \frac{\$1{,}000}{\left(1 + \frac{2.5754\%}{2}\right)^2} \tag{7.16}$$

$$\$984.430 = \frac{\$1{,}000}{\left(1 + \frac{1.5754\%}{2}\right)^2} \tag{7.17}$$

The fact that the possible values of the one-year spot rate can be extracted from the tree is at first surprising. The starting point of the example is the date 0 values of the 0.5-, 1-, and 1.5-year spot rates, along with assumptions about the evolution of the six-month rate over the next years. But because this information, in combination with arbitrage or risk-neutral arguments, is sufficient to determine the price tree of the 1.5-year zero, it is also sufficient to determine the possible values of the one-year spot rate in six months. Put another way, having specified initial

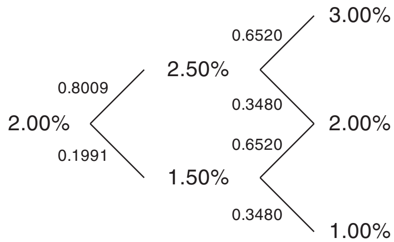

FIGURE 7.6 Risk-Neutral Process for the Six-Month Rate.

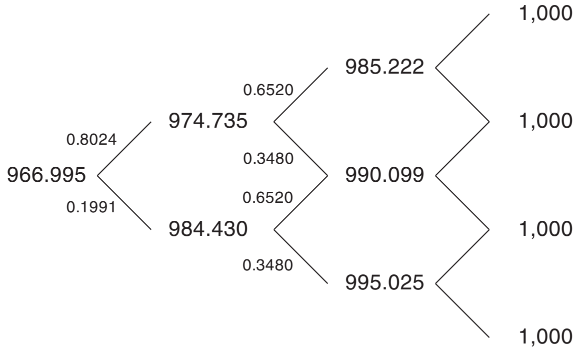

FIGURE 7.7 Final Price Tree for a 1.5-Year Zero Coupon Bond, with Probabilities.

spot rates and the evolution of the six-month rate, a modeler may not make any further assumptions about the behavior of the one-year rate.

The six-month rate process completely determines the one-year rate process here, because the model presented has only one factor. Writing down a tree for the evolution of the six-month rate alone implicitly assumes that prices of all fixed income securities can be determined by the evolution of that rate. A multi-factor term structure model is presented in Chapter 9.

This section concludes with two additional observations about multi-period settings. First, extending the tree to any number of dates requires assumptions about the future possible values of the short-term rate and the calculation of risk-neutral probabilities that recover a given set of bond prices. Second, the composition of replicating portfolios depends on date and state. For example, the replicating portfolio of a derivative as of date 0 is usually different from its replicating portfolio on date 1, state 0, and different again from its replicating portfolio on date 1, state 1. From a trading perspective, this means that replicating portfolios must be adjusted as time passes and as interest rates change. These adjustments are known as *dynamic replication*, in contrast to the *static replication strategies* of earlier chapters, like replicating a coupon bond with an unchanging portfolio of two other coupon bonds of the same maturity.

7.5 PRICING A CONSTANT-MATURITY TREASURY SWAP

Equipped with the tree in Figure 7.7, this section prices a \$1,000,000 stylized *constant-maturity Treasury (CMT) swap* struck at 2%. This swap pays,

$$\$1{,}000{,}000\frac{y_{CMT} - 2\%}{2} \tag{7.18}$$

every six months until it matures, where y_{CMT} is the semiannually compounded yield – of a predetermined maturity – on the payment date. This example prices a one-year CMT swap on the six-month yield, though, in practice, CMT swaps trade most commonly on the yields of the most liquid bonds, for example, on two-, five- and 10-year Treasury yields.

Because six-month semiannually compounded yields equal six-month spot rates, rates from the tree of the previous section can be substituted into Equation (7.18) to calculate the payoffs of the CMT swap. On date 1, the state 1 and state 0 payoffs are, respectively,

$$\$1{,}000{,}000\frac{2.50\% - 2\%}{2} = \$2{,}500 \tag{7.19}$$

$$\$1{,}000{,}000\frac{1.50\% - 2\%}{2} = -\$2{,}500 \tag{7.20}$$

Similarly on date 2, the state 2, 1, and 0 payoffs are, respectively,

$$\$1{,}000{,}000\frac{3\% - 2\%}{2} = \$5{,}000 \tag{7.21}$$

$$\$1{,}000{,}000\frac{2\% - 2\%}{2} = \$0 \tag{7.22}$$

$$\$1{,}000{,}000\frac{1\% - 2\%}{2} = -\$5{,}000 \tag{7.23}$$

The possible values of the CMT swap at maturity, on date 2, are given by Equations (7.21) through (7.23). The possible values on date 1 are given by the expected discounted value of the date 2 payoffs under the risk-neutral probabilities plus the date 1 payoffs given by (7.19) and (7.20). The resulting date 1 values in states 1 and 0 are, respectively,

$$\frac{.6520 \times \$5{,}000 + .3480 \times \$0}{1 + \frac{.0250}{2}} + \$2{,}500 = \$5{,}719.52 \tag{7.24}$$

$$\frac{.6520 \times 0 + .3480 \times (-\$5{,}000)}{1 + \frac{.0150}{2}} - \$2{,}500 = -\$4{,}227.29 \tag{7.25}$$

Finally, the value of the swap on date 0 is the expected discounted value, under the risk-neutral probabilities, of the date-1 payoffs, given by Equations (7.24) and (7.25),

$$\frac{.8009 \times \$5{,}719.52 + .1991 \times (-\$4{,}227.29)}{1 + \frac{.0200}{2}} = \$3{,}702.11 \tag{7.26}$$

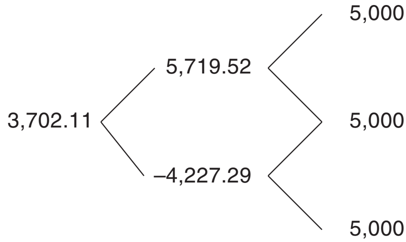

FIGURE 7.8 Price Tree for a Stylized CMT Swap.

The tree in Figure 7.8 summarizes the value of the stylized CMT swap over dates and states. A value of \$3,702.11 for the CMT swap might seem surprising at first. After all, the cash flows of the CMT swap are zero at a rate of 2%, and 2% is, under the true probabilities, the average rate on each date. The explanation, of course, is that the risk-neutral probabilities, not the true probabilities, determine the arbitrage price of the swap. The expected discounted value of the swap under the true probabilities can be computed by following the steps leading to Equations (7.24) through (7.26) but using the probability 0.5 for all up- and down-moves. The result of these calculations does give a value close to zero, namely, −\$6.07.

7.6 OPTION-ADJUSTED SPREAD

Option-adjusted spread is a widely used measure of the relative value of a security, that is, of its market price relative to its model value. OAS is defined as the spread such that the market price of a security is recovered when that spread is added to discount rates in the model. To illustrate, say that the market price of the CMT swap in the previous section is \$3,699.18, \$2.92 less than the model price. In that case, the OAS of the CMT swap turns out to be 10 basis points. To see this, add 10 basis points to the discounting rates of 2.5% and 1.5% in Equations (7.24) and (7.25), respectively, to get new swap values of,

$$\frac{.6520 \times \$5{,}000 + .3480 \times \$0}{1 + \frac{.0260}{2}} + \$2{,}500 = \$5{,}717.93 \tag{7.27}$$

$$\frac{.6520 \times 0 + .3480 \times (-\$5{,}000)}{1 + \frac{.0160}{2}} - \$2{,}500 = -\$4{,}226.43 \tag{7.28}$$

Note that, when calculating value with an OAS spread, rates are only shifted for the purpose of discounting. Rates are not shifted for the purposes of computing cash flows. In the CMT swap example, cash flows are still computed using Equations (7.19) through (7.23).

Completing the valuation with an OAS of 10 basis points, use the results of (7.27) and (7.28) and a discount rate of 2% plus the OAS spread of 10 basis points, or 2.10%, to obtain an initial CMT swap value of,

$$\frac{.8009 \times \$5{,}717.93 + .1991 \times (-\$4{,}226.43)}{1 + \frac{.0210}{2}} = \$3{,}699.18 \tag{7.29}$$

Hence, as claimed, discounting at the risk-neutral rates plus an OAS of 10 basis points in the model recovers the given market price of \$3,699.18. If a security's OAS is positive, its market price is less than its model price, which means that the security trades cheap. If the OAS is negative, the security trades rich.

Another perspective on the relative value implications of an OAS spread is the fact that the expected return of a security with an OAS, under the risk-neutral process, is the short-term rate plus the OAS per period. Very simply, discounting a security's expected value by a particular rate per period is equivalent to that security's earning that rate per period. In the example of the CMT swap, the expected return of the fairly priced swap under the risk-neutral process over the six months from date 0 to date 1 is,

$$\frac{.8009 \times \$5{,}719.52 - .1991 \times \$4{,}227.29 - \$3{,}702.11}{\$3{,}702.11} = 1.00\% \tag{7.30}$$

which is six months' worth of the initial rate of 2%. On the other hand, with an OAS of 10 basis points, the expected return of the cheap swap is,

$$\frac{.8009 \times \$5{,}717.93 - .1991 \times \$4{,}226.43 - \$3{,}699.18}{\$3{,}699.18} = 1.05\% \tag{7.31}$$

which is six months' worth of the initial rate of 2% plus the OAS of 10 basis points, or half of 2.10%.

7.7 PROFIT AND LOSS ATTRIBUTION WITH AN OAS

Chapter 3 introduced profit and loss (P&L) attribution. This section gives a mathematical description of attribution in the context of term structure models and of securities that trade with an OAS. While the notation of this chapter is quite formal, the presentation remains intuitive.

By the definition of a one-factor model, and by the definition of OAS, the market price of a security at time t and a factor, r, which is often a rate, can be written as $P_t(r, OAS)$. Using a first-order Taylor approximation, the change in the price of the security is,

$$dP = \frac{\partial P}{\partial r}dr + \frac{\partial P}{\partial t}dt + \frac{\partial P}{\partial OAS}dOAS \tag{7.32}$$

where $\partial P/\partial r$ gives the change in the price of the security for a change in r, holding t and OAS constant; $\partial P/\partial t$ gives the change in price for a change in t holding r and OAS constant; and the same for $\partial P/\partial OAS$. In words, Equation (7.32) breaks down the total change in price to components of change due to changes in r, t, and OAS.

Dividing both sides of Equation (7.32) by price and taking expectations,

$$E\left[\frac{dP}{P}\right] = \frac{1}{P}\frac{\partial P}{\partial r}E[dr] + \frac{1}{P}\frac{\partial P}{\partial t}dt \tag{7.33}$$

Note that dP/P is the change in price divided by price, or the percentage change in price. Because the OAS calculation assumes that OAS is constant over the life of the security, moving from (7.32) to (7.33) assumes that the expected change in the OAS is zero.

As mentioned in the previous section, if expectations are taken with respect to the risk-neutral process, then, for any security priced according to the model,

$$E\left[\frac{dP}{P}\right] = r_0 dt \tag{7.34}$$

But Equation (7.34) does not apply to securities that are not priced according to the model, that is, to securities with an OAS not equal to zero. For these securities, by definition, the cash flows are discounted not at the short-term rate, but at the short-term rate plus the OAS. Equivalently, as argued in the previous section, the expected return under the risk-neutral probabilities is not the short-term rate, but the short-term rate plus the OAS. Hence, the more general form of Equation (7.34) is,

$$E\left[\frac{dP}{P}\right] = (r_0 + OAS)dt \tag{7.35}$$

Combining these pieces, substitute for $(1/P)\partial P/\partial t$ from (7.33) and then for $E[dP/P]$ from (7.35) into Equation (7.32) and rearrange terms, which breaks down the return of a security into its component parts,

$$\frac{dP}{P} = (r_0 + OAS)dt + \frac{1}{P}\frac{\partial P}{\partial r}(dr - E[dr]) + \frac{1}{P}\frac{\partial P}{\partial OAS}dOAS \tag{7.36}$$

Finally, multiplying through by P,

$$dP = (r_0 + OAS)Pdt + \frac{\partial P}{\partial r}(dr - E[dr]) + \frac{\partial P}{\partial OAS}dOAS \tag{7.37}$$

In words, the return of a security or its P&L may be divided into a component due to the passage of time; a component due to changes in the factor; and a component due to the change in the OAS. In the language of Chapter 3, the terms on the right-hand side of (7.37) represent, in order, carry–roll-down, gains or losses from rate changes, and gains or losses from spread change.[1] For models with predictive power, the OAS converges or trends to zero; that is, the security price converges or trends toward its fair value according to the model.

The decompositions of Equations (7.36) and (7.37) highlight the usefulness of OAS as a measure of value. If a model is correct, a long position in a cheap security earns superior returns in two ways. First, it earns the OAS over time intervals in which the security does not converge to its fair value. Second, it earns its sensitivity to OAS times any convergence of that OAS to zero.

The decompositions also provide a framework for relative value trading. When a cheap or rich security is identified, a relative value trader buys or sells the security and hedges out all interest rate or factor risk. Mathematically, $\partial P/\partial r = 0$. In that case, the expected return or P&L depends only on the short-term rate, the OAS of the securities traded, and any OAS convergence. If the trader finances the trade at the short-term rate, that is, borrows P at rate r_0 to purchase the security, then the expected return is simply equal to the OAS plus any convergence return. If the hedge itself costs or generates funds, then the P&L also includes a return on those funds at the short-term rate. If the hedging securities are not fairly priced relative to the model, but have an OAS of their own, then the P&L also includes an OAS on the hedge. Finally, Chapter 8 explains that bearing interest rate or factor risk may earn a risk premium, in which case there would be an additional term in Equations (7.34) through (7.37) that depends on the amount of factor risk borne. But, in the relative value context, where factor risk is hedged away, any risk premium terms cancel out, and the P&L of the trade is as described in this paragraph.

7.8 REDUCING THE TIME STEP

This chapter has so far assumed that the time elapsed between dates of the tree is six months. The methodology outlined, however, adapts easily to any

[1] For expositional simplicity no explicit coupon or other direct cash flows have been included in this discussion.

time step of Δt years. For monthly time steps, for example, $\Delta t = 1/12$ or .0833, and one-month rather than six-month interest rates appear on the tree. Furthermore, discounting is done over the appropriate time interval. If the rate of term Δt is r, then discounting means dividing by $1 + r\Delta t$. In the case of monthly time steps, discounting with a one-month rate of 2% means dividing by $1 + 0.02/12$.

In practice there are two reasons to choose time steps smaller than six months. First, a security or portfolio of securities rarely makes all of its payments in even six-month intervals from the starting date. Reducing the time step to a month, a week, or even a day can ensure that all cash flows are sufficiently close in time to some date in the tree. Second, assuming that the six-month rate can take on only two values in six months, three values in one year, and so on, produces a tree that is too coarse for many practical pricing problems. Reducing the step size can fill the tree with enough rates to price contingent claims with sufficient accuracy.

While smaller time steps generate more realistic interest rate distributions, they require that more attention be paid to numerical issues, and they may make computations too slow for their intended uses. The choice of step size ultimately depends, therefore, on the problem at hand. When pricing a 30-year callable bond, for example, a model with weekly or even monthly time steps may provide a realistic enough interest rate distribution to generate reliable prices. By contrast, pricing a one-month bond option with any precision would require a much smaller time step. While the trees in this chapter assume that the step size is the same throughout the tree, this need not be the case. Sophisticated implementations of trees allow step size to vary across dates in order to achieve a balance between realism and computational concerns.

7.9 FIXED INCOME VERSUS EQUITY DERIVATIVES

The famous Black-Scholes-Merton (BSM) pricing analysis of stock options can be summarized as follows. Under the assumptions that the stock price evolves according to a particular random process and that the short-term interest rate is constant, it is possible to form a portfolio of the underlying stock and short-term bonds that replicates the option. Therefore, by arbitrage arguments, the price of the option equals the price of the replicating portfolio.

Consider applying this logic to price an option on a five-year bond. The starting point might be an assumption about how the price of the five-year bond evolves over time, but the task is considerably more complicated than for the price of a stock. First, the price of a bond converges to its face value at maturity, while the stock price has no similar constraint. Second, because of the maturity constraint, the volatility of a bond's price must decline as the bond approaches maturity. Hence, the simple assumption that the volatility of a stock is constant is not as appropriate for bonds. Third, because stock volatility is very large relative to short-term rate volatility, it is often acceptable to assume that the short-term interest rate is constant. By contrast, it seems inconsistent to assume that bond prices – which depend on interest rates – follow some random process, while assuming that the short-term interest rate is constant.

These objections led researchers to make assumptions about the random evolution of the interest rate rather the bond price. In that way, bond prices naturally approach par, price volatilities naturally approach zero, and no interest rate is assumed to be constant. But this approach raises another set of questions. Which interest rate is assumed to evolve in a particular way? Assumptions about the five-year spot rate, for example, are not sufficient for two reasons. First, five-year coupon bond prices depend on shorter-term spot rates as well. Second, an option on a particular five-year bond soon depends on the prices of a bond that is no longer a five-year bond, but a bond with slightly less time to maturity. Therefore, assumptions usually have to be made about the evolution of the entire term structure of interest rates to price bond options and other derivatives. This chapter shows that, in a one-factor model, assumptions about the evolution of the short-term rate are sufficient to model the evolution of the entire term structure.

In short, there are several arguments to move beyond BSM in the fixed income context. Nevertheless, for simplicity, practitioners do use versions of BSM to price and hedge several classes of fixed income derivatives. These methodologies are described at length in Chapter 16.

CHAPTER 8

Expectations, Risk Premium, Convexity, and the Shape of the Term Structure

Chapter 7 shows how bonds and other interest rate contingent claims can be priced given the evolution of the short-term rate. This chapter shows how the shape of the term structure of interest rates is determined by assumptions about the evolution of the short-term rate and by assumptions about the risk premium demanded by investors for bearing interest rate risk. The first few sections of the chapter present concepts by way of example, in the simple, binomial tree framework of the previous chapter. The last section of the chapter presents the same ideas in more general setting, though at the cost of some higher-level mathematics. Chapter 9 concludes the presentation of term structure models with a detailed description of two well-known models.

8.1 EXPECTATIONS

Consider a simple framework with annual periods. Assume for the moment that the current one-year rate is 8%, and that investors know with certainty that the one-year rate in one year will be 7% and in two years will be 6%. Then, the prices of one-, two-, and three-year zero coupon bonds with a unit face value, $P(1)$, $P(2)$, and $P(3)$, are priced such that,

$$P(1) = \frac{1}{1.08}$$

$$P(2) = \frac{1}{1.08 \times 1.07}$$

$$P(3) = \frac{1}{1.08 \times 1.07 \times 1.06} \tag{8.1}$$

But by the definition of forward rates (see Chapter 2), Equations (8.1) say that the first three forward rates are 8%, 7%, and 6%. Hence, with investor certainty as to future interest rates, that is, without any volatility around those expectations, the term structure of interest rates – here expressed in terms of forward rates – is completely determined by expectations. Consequently, depending on expectations, the term structure can take on any shape: flat, upward-sloping, downward-sloping, or some combination of these.

In practice, expectations cannot sensibly take on any arbitrary pattern. The financial community can have very specific views about short-term rates over short horizons, derived, for example, from anticipation of policy changes on central bank meeting dates and from the supply and demand conditions for funds (e.g., tax payment dates, the bond issuance calendar, quarterly balance sheet management). Over longer horizons, however, expectations are not as granular. Analysis of money market conditions is unlikely to reveal, for example, that the expected one-year rate in 29 years is very different from the expected one-year rate in 30 years. On the other hand, macroeconomic analysis might argue that the long-run expectation of the short-term rate is 4%: 1% due to the long-run real-rate of interest and 3% due to long-run inflation.

8.2 VOLATILITY AND CONVEXITY

While investors have expectations about future short-term rates, they recognize the limits of their analyses, that is, realized rates are assumed to fluctuate randomly around expectations. Continuing with the framework of the previous chapter, consider the binomial tree for the one-year rate in the top part of Figure 8.1. The step size is one year, and the probabilities of all transitions are 50% (not shown). The level of rates and their volatility is exaggerated in this tree to illustrate the concepts of this chapter. Note that the expected value of the short-term rate in one year is 9%, as is the expected short-term rate in two years,

$$50\% \times 13\% + 50\% \times 5\% = 9\% \quad (8.2)$$

$$50\%[50\% \times 17\% + 50\% \times 9\%] + 50\%[50\% \times 9\% + 50\% \times 1\%] = 9\% \quad (8.3)$$

Note also that the volatility of the change in rate at any transition is 4%, or 400 basis points. For example, with the mean of the first transition calculated in Equation (8.2) to be 9%, the volatility of that transition is,

$$\sqrt{50\%[13\% - 9\%]^2 + 50\%[5\% - 9\%]^2} = 4\% \quad (8.4)$$

The price of a one-year zero in this model is simply 1/1.09, or 0.917431. Assuming, for the moment, that investors are risk neutral, the price trees of the two- and three-year zeros can be calculated by expected discounted

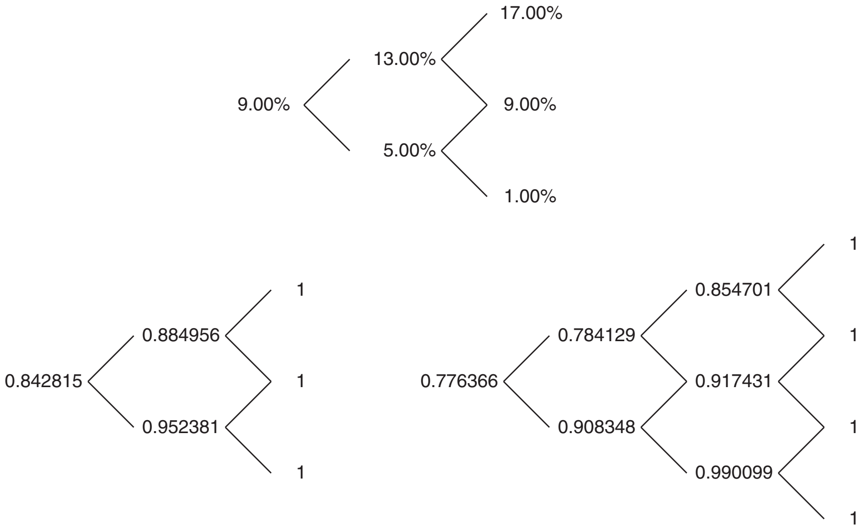

FIGURE 8.1 Binomial Rate Tree and Price Trees for Two- and Three-Year Zero Coupon Bonds. Steps Are Annual, and the Probabilities of All Transitions Are 50%.

TABLE 8.1 Prices of Zero Coupon Bonds and Their Associated Forward Rates from the Rate Tree in Figure 8.1. Rates Are in Percent.

Term	Price	Forward Rate
1	0.917431	9.0000
2	0.842815	8.8532
3	0.776366	8.5590

value, as explained in the previous chapter. These trees are shown in the bottom of Figure 8.1, with the supporting calculations left to the reader. Table 8.1 collects the prices of the three zero coupon bonds, along with the associated forward rates. The striking feature of the table is that the term structure of forward rates is downward sloping, despite the fact that interest rate expectations are flat at 9%. This result can be explained by the interaction of interest rate volatility with the convexity of bond prices.

A short detour is required at this point to present and explain *Jensen's inequality* as applied to bond pricing. For a random variable, like the one-year rate, r,

$$E\left[\frac{1}{1+r}\right] > \frac{1}{E[1+r]} = \frac{1}{1+E[r]} \tag{8.5}$$

In words, the expected price of a bond is greater than the price of a bond at the expected interest rate.

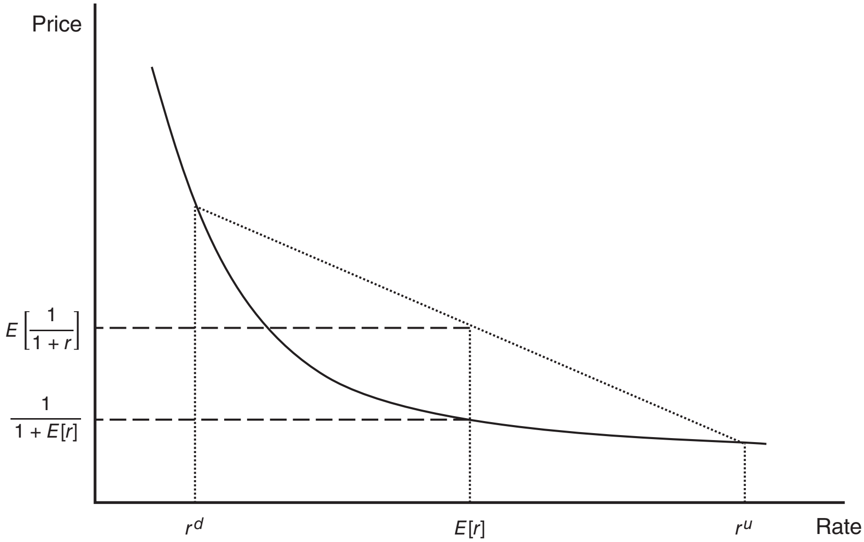

FIGURE 8.2 An Illustration of Jensen's Inequality as Applied to Bond Pricing.

This inequality is easily explained by Figure 8.2. In the figure, the rate can take on two values, r^d and r^u, with equal probability, resulting in an expected value just between them, $E[r]$. Each possible value of r has an associated price, and the expected value of that price, $E[1/(1+r)]$, is graphically depicted as the vertical-axis coordinate of the dotted line connecting the points $\{r^d, 1/(1+r^d)\}$ and $\{r^u, 1/(1+r^u)\}$. Because of the curvature or convexity of the price–rate curve, however, this expected price exceeds the price at a rate of $E[r]$, which is $1/(1+E[r])$. And this is exactly the relationship described in Equation (8.5).

Returning to the role of volatility and convexity, let f denote the one-year rate, one year forward, and consider the date-0 price of a two-year zero coupon bond, as expressed in Equation (8.6). By definition, the price of the two-year zero equals its unit face amount discounted by 9% over the first year and by f over the second year. By the logic of pricing along the tree, this price also equals the discounted expected value of the date-1 price of the bond. Multiplying both sides of (8.6) by 1.09 and invoking Jensen's inequality in Equation (8.5), gives (8.7). And from this equation, (8.8) follows directly: the one-year rate, one year forward, is less than the expected one-year rate in one year,

$$0.8428 \equiv \frac{1}{(1.09)(1+f)} = \frac{1}{1.09}\left[50\% \times \frac{1}{1.13} + 50\% \times \frac{1}{1.05}\right] \tag{8.6}$$

$$\frac{1}{1+f} > \frac{1}{50\% \times 1.13 + 50\% \times 1.05} = \frac{1}{1.09} \tag{8.7}$$

$$f < 9\% \tag{8.8}$$

8.3 AN ANALYTICAL DECOMPOSITION OF FORWARD RATES

This section derives a general decomposition of forward rates in terms of expectations, convexity, and risk premium. The level of mathematics here is higher than used in most of the book, but the discussion still aims at intuition.

Assume that all bond prices are determined by the instantaneous rate, r, which takes on the value of r_t at time t. Let $P_t(r_t, T)$ be the price of a T-year zero coupon bond at time t. By *Ito's lemma*, a discussion of which is beyond the scope of this book,

$$dP = \frac{\partial P}{\partial r}dr + \frac{\partial P}{\partial t}dt + \frac{1}{2}\frac{\partial^2 P}{\partial r^2}\sigma^2 dt \tag{8.9}$$

where dP, dr, and dt are the changes in price, rate, and time over the next instant, respectively, and σ is the volatility of changes in r. The two first-order partial derivatives in Equation (8.9) denote the instantaneous change in the bond price for a unit change in the rate (with time unchanged) and for a unit change in time (with rate unchanged), respectively. Finally, the second order partial derivative in the equation gives the instantaneous change in $\partial P/\partial r$ (with time unchanged). Dividing both sides of (8.9) by price,

$$\frac{dP}{P} = \frac{1}{P}\frac{\partial P}{\partial r}dr + \frac{1}{P}\frac{\partial P}{\partial t}dt + \frac{1}{2}\frac{1}{P}\frac{\partial^2 P}{\partial r^2}\sigma^2 dt \tag{8.10}$$

Equation (8.10) breaks down the instantaneous return on the zero coupon bond into three components, but this decomposition can be written more intuitively by invoking several ideas from earlier chapters.

First, in terms of instantaneous compounded forward rates, $f(t)$, the price of a T-year zero coupon bond is (from Section A2.1),

$$P = e^{-\int_0^T f(s)ds} \tag{8.11}$$

Then, differentiating both sides of (8.11) with respect to t, recognizing that increasing t decreases T one-for-one,

$$\frac{\partial P}{\partial t} = -\frac{\partial P}{\partial T} = f(T)P \tag{8.12}$$

Second, by the definitions of duration, D, and convexity, C,

$$D \equiv -\frac{1}{P}\frac{\partial P}{\partial r} \tag{8.13}$$

$$C \equiv \frac{1}{P}\frac{\partial^2 P}{\partial r^2} \tag{8.14}$$

Now, substituting Equations (8.12) through (8.14) into the return decomposition (8.10),

$$\frac{\partial P}{P} = f(T)dt - Ddr + \frac{1}{2}C\sigma^2 dt \tag{8.15}$$

Equation (8.15) gives the return decomposition in terms of the following three components. The first is the return due to the passage of time, which, in this case, is the forward rate, $f(T)$.[1] The second and third components are returns due to changes in the rate. The second term says that increases in rate reduce bond return in proportion to duration. The third term says that the volatility of rates – movement of rates either up or down – increases return in proportion to convexity. To appreciate this term, recall from Chapter 4 that, across portfolios with the same duration, more convex portfolios increase more in value as rates change (at a fixed moment in time), whether rates rise or fall.

To draw conclusions about expected returns, take the expectation of both sides of (8.15),

$$E\left[\frac{\partial P}{P}\right] = f(T)dt - DE[dr] + \frac{1}{2}C\sigma^2 dt \tag{8.16}$$

The intuition of this decomposition is the same as for Equation (8.15), but with the duration component depending not on the change in rate but on the expected change in rate.

The next step in the analysis introduces the concept of a risk premium. Risk-neutral investors, who do not require a risk premium, demand that each bond offer an expected return equal to the short-term rate of interest. Mathematically,

$$E\left[\frac{dP}{P}\right] = r_0 dt \tag{8.17}$$

Risk averse investors, however, demand higher expected returns for bonds with greater interest rate risk. The appendix to this chapter shows that the interest rate risk of a bond over the next instant may be measured by its duration with respect to the interest rate factor, and that risk-averse investors demand a risk premium proportional to duration. This risk premium may depend on time and on the level of rates, but not on the characteristics of any individual bond. The discussion proceeds here,

[1] Note that the result here is related to the result in Chapter 3 that the T-year zero coupon bond earns the forward rate corresponding to its term, $f(T)$, under the assumption of an unchanged term structure and, implicitly, no change in the short-term rate and no interest rate volatility.

however, as if the risk premium were constant and denoted by λ. In that case, the expected return equation for risk-averse investors is,

$$E\left[\frac{dP}{P}\right] = r_0 dt + \lambda D dt \tag{8.18}$$

Say, for example, that the short-term rate is 1%, that the duration of a bond is five, and that the risk premium is 10 basis points per year of duration risk. Then, according to Equation (8.18), the bond's expected return is 1% + $5 \times 0.1\% = 1.5\%$ per year.

Another useful way to think of the risk premium is in terms of the *Sharpe ratio* (SR) of a security, defined as its expected excess return (over the short-term rate) divided by the standard deviation of its return. Because the random part of a bond's return comes from its duration times the change in rate, as in Equation (8.15), the standard deviation of the return equals the duration times the standard deviation of rates. Therefore, the SR of a bond may be written as,

$$SR = \frac{E[dP/P] - r_0 dt}{\sigma D dt} = \frac{\lambda}{\sigma} \tag{8.19}$$

where the second equality follows from Equation (8.18). For example, if the risk premium is 10 basis points per year, and if the standard deviation of rates is 100 basis points per year, then the Sharpe ratio of bond investments is 10%.

The decomposition of returns can now be combined with the economics of the risk premium to draw conclusions about the shape of the term structure of forward rates. Equating the expressions for expected returns in the right-hand sides of Equations (8.16) and (8.18),

$$f(T) = \left\{ r_0 + E\left[\frac{dr}{dt}\right] D \right\} + \lambda D - \frac{1}{2} C \sigma^2 \tag{8.20}$$

Equation (8.20) mathematically describes the determinants of forward rates. The three terms represent the impacts of expectations, risk premium, and convexity, respectively. The first term says that the forward rate is composed of the instantaneous interest rate plus the expected change in that rate times the duration of the zero coupon bond corresponding to the term of the forward rate. In other words, the higher the instantaneous rate, the higher the forward rate; the more rates are expected to increase, the higher the forward rate; and the greater the corresponding duration, the greater the effect of expected rate changes on the forward rate.

The second term on the right-hand side of (8.20) says that the forward rate increases with the risk premium in proportion to the corresponding duration. In other words, the greater the corresponding interest rate risk and the greater the risk premium, the greater the forward rate.

Chapter 7 noted that a drift in the short-term rate of a certain number of basis points has the same effect on bond pricing as a risk premium of that number of basis points per year of duration risk. Equation (8.20) formalizes this statement. Increasing the risk premium or increasing the expected short-term rate by the same amount are indistinguishable from the observation of forward rates. This means that the term structure of interest rates cannot, on its own, be used to separate expectations of rate changes from risk premium. From a modeling perspective, this means that only the risk-neutral process is relevant for pricing. Dividing the risk-neutral drift into expectations and risk premium might be very useful for economic perspectives and for macro-style trading (see Chapter 9), but this division is not observable from a cross section of bond prices alone.

The first two terms of Equation (8.20) can also be cast in terms of theories of the term structure of interest rates. (Put aside the convexity term for the moment.) Under the *pure expectations hypothesis*, the risk premium, λ, is zero, and the term structure of forward rates is determined by expectations, $E[dr/dt]$. In this view of the world, the most natural "no-change" scenario, in the terms of Chapter 3, is that short-term rates evolve as expected and that forward rates are realized. At the opposite extreme, under the *pure risk premium hypothesis*, the market has no expectations about rates, that is, $E[dr/dt] = 0$, and the term structure of forward rates is determined by the risk premium. In this view of the world, the most natural "no-change" scenario is that short-term rates stay the same, as expected, which, in terms of Chapter 3 is an unchanged term structure. The reality, of course, can be between the two extremes, such the the term structure is determined by a mix of expectations and risk premium.

To conclude the discussion of Equation (8.20), the third term shows that the forward rate is reduced because of volatility and the convexity of the zero corresponding to the term of the forward rate by $0.5C\sigma^2$. Using this to reinterpret Equation (8.16), the indirect reduction in return through the forward rate, because of convexity, is exactly offset by the direct increase in return, because of convexity. Put another way, the expected return condition of Equation (8.18) ensures that there is no net advantage of convexity. The significance of this reasoning for investment and hedging decisions is introduced in Chapter 4 in the context of establishing long- and short-convexity positions.

CHAPTER 9

The Vasicek and Gauss+ Models

This chapter, the last of the three on term structure models, presents the well-known Vasicek and Gauss+ models. The Vasicek model started the literature on short-term rate models;[1] remains an extremely good starting point for learning about these models; and can still be used in some applied contexts. The Gauss+ model has proven very popular for proprietary trading, for both relative value and macro-style trading. The presentation of this model here is directed toward determined readers who would like to implement a term structure model for their own trading purposes.

9.1 THE VASICEK MODEL

The Vasicek model assumes *mean reversion* to set the expected path of the short-term rate. When below its long-term value, the short-term rate is expected to increase; when above its long-term value, the short-term rate is expected to decrease. Mathematically, the risk-neutral dynamics for the short-term rate, r, are given by,

$$dr = k(\theta - r)dt + \sigma dw \tag{9.1}$$

In words, the instantaneous change in the short-term rate, dr, is determined by a trend or drift plus a random fluctuation or shock. The drift is equal to the parameter of mean reversion, k, times the distance between the long-run value of the short-term rate, θ, and its current value. Because all variables are expressed in annualized terms, the dt factor adjusts for the actual passage of time. Say, for example, that $r = 2\%$; $k = 0.0165$, and $\theta = 11\%$. Then the drift of the short-term rate in Equation (9.1) is $0.0165(11\% - 2\%)$ or 0.1485% or 14.85 basis points per year. The drift over a month, therefore, with $dt = 1/12$, would be 14.85/12 or about 1.2 basis points per month. The shock around the drift in Equation (9.1) is σdw, where dw is a normally distributed random variable with mean equal to zero and

[1] Vasicek, O. (1977), "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics* 5.

standard deviation equal to $\sqrt{dt}$. The shock, therefore, is normally distributed with mean zero and standard deviation $\sigma\sqrt{dt}$. For example, if σ is 0.95%, or 95 basis points per year, then the volatility of the shock over a month is $95 \times \sqrt{1/12} = 27.4$ basis points.

As explained in the previous chapter, fixed income security prices may incorporate a risk premium that is indistinguishable from a drift in the evolution of the short-term rate. Along these lines, Equation (9.1) can be viewed as containing a drift due to a risk premium. Assume for the purposes of this section that the risk premium is a known constant of λ basis points per year, and that the long-run value of the short-term rate under the true or real-world probabilities is r_∞. In that case, the true process of the short-term rate with the addition of a drift due to the risk premium is,

$$\begin{aligned} dr &= k(r_\infty - r)dt + \lambda dt + \sigma dw \\ &= k\left(\left[r_\infty + \frac{\lambda}{k}\right] - r\right)dt + \sigma dw \end{aligned} \tag{9.2}$$

$$\theta \equiv r_\infty + \frac{\lambda}{k} \tag{9.3}$$

Equation (9.3) neatly emphasizes the inability to distinguish expectations from risk premium by observing security prices: an infinite number of combinations of r_∞ and λ give the same θ and, therefore, the same risk-neutral price process in Equation (9.1).

One reason that the Vasicek model is useful, both for learning about term structure models and for some simple pricing and hedging applications, is that most rates and prices from the model can be expressed through simple formulae. For the most complex securities, numerical methods, like binomial trees, are needed, and Appendix A9.1 explains how the model's dynamics might be captured in a binomial tree. The text, however, continues by presenting analytic solutions of the model, of which some of the most useful are,

$$E[r_t] = r_0 e^{-kt} + \theta(1 - e^{-kt}) \tag{9.4}$$

$$V[r_t] = \sigma^2 \frac{1 - e^{-2kt}}{2k} \tag{9.5}$$

$$f(t) = \theta + e^{-kt}(r_0 - \theta) - \frac{\sigma^2}{2k^2}(1 + e^{-2kt} - 2e^{-kt}) \tag{9.6}$$

$$\hat{r}(t) = \theta + \frac{1 - e^{-kt}}{kt}(r_0 - \theta) - \frac{\sigma^2}{2k^2}\left(1 + \frac{1 - e^{-2kt}}{2kt} - 2\frac{1 - e^{-kt}}{kt}\right) \tag{9.7}$$

where $E[r_t]$ gives today's expectation of the short-term rate at time t, $V[r_t]$ gives the variance of the short-term rate at time t, $f(t)$ is the continuously compounded forward rate of term t; and $\hat{r}(t)$ is the continuously compounded spot rate of term t.

Figures 9.1 through 9.3 illustrate these formulae with the parameter values given earlier. The expected short-term rate, according to Equation (9.4) and the solid line in Figure 9.1, moves gradually from r_0 today ($t = 0$) to θ in the very distant future ($t = \infty$). The mean reversion parameter governing the speed of that adjustment, $k = 0.0165$, is sometimes quoted instead as a *half-life*. From Equation (9.4), a shock to r_0 decays according to the factor e^{-kt}. And half of such a shock decays away after time h, such that,

$$e^{-kh} = \frac{1}{2}$$
$$h = ln(2)/k \tag{9.8}$$

For the relatively small mean reverting parameter, $k = 0.0165$, h is over 42 years, which means that any shock to r affects rate expectations over a very long period of time. Equivalently, the expected rate takes a very long time to revert from the current rate to θ.

The standard deviation of the short-term rate around its expectations is given by the square root of (9.5) and shown by the dashed lines in Figure 9.1. With a volatility parameter of 95 basis points per year and a very slow mean reversion, the standard deviation around expectations is quite wide. The figure does show, however, that mean reversion narrows this standard deviation. Without mean reversion, the standard deviation of the short-term rate is greater, at $\sigma\sqrt{t}$. Put another way, the pull of the short-term rate to the constant value θ reduces the standard deviation of the short-term rate as of any future date.

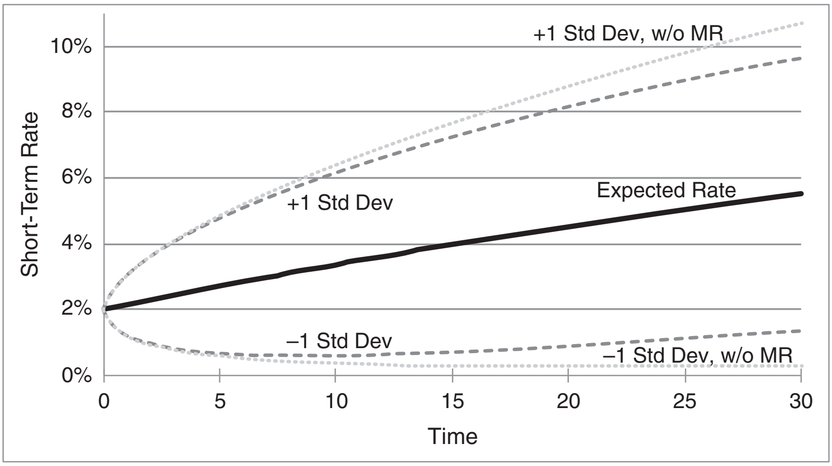

FIGURE 9.1 Expectations of the Continuously Compounded Short-Term Rate in the Vasicek Model, with One-Standard-Deviation Bands. Light Dotted Lines Give Bands Without Any Mean Reversion. Model Parameters Are $r_0 = 2\%$, $\theta = 11\%$, $k = 0.0165$, and $\sigma = 0.95\%$.

Equations (9.6) and (9.7), illustrated in Figure 9.2, give the continuously compounded forward and spot rates in the model. The shape of the forward curve is discussed presently. Recall, however, as discussed in Chapter 2, so long as the forward curve is above the spot curve, spot rates are increasing.

Figure 9.3 shows the term structure of forward rate volatility in the model, that is, the instantaneous volatility of forward rates of different terms. Because the only volatility in the model is the volatility of changes in the short-term rate, the volatility of $f(t)$ – from Equation (9.6) – is just σe^{-kt}. The mean reversion feature of the model, therefore, captures the empirical regularity that, for longer terms, the term structure of volatility is downward sloping. Empirical volatilities of short-term rates are much lower than indicated in this figure, however, because the central bank pegs short-term rates. The Gauss+ model, discussed next, has the flexibility to capture both a low short-term rate volatility and an ultimately declining term structure of volatilities. In any case, note from Equation (9.6) that the sensitivity of each forward rate to changes in the short-term rate is e^{-kt}. Hence, these sensitivities across terms have the same shape as in Figure 9.3.

Figure 9.4, the last presented on the Vasicek model, decomposes the forward rate curve into expectations, risk premium, and convexity using the values $\lambda = 0.125\%$, which – given $\theta = 11\%$ and Equation (9.3) – means that $r_\infty = 3.424\%$. In this decomposition, expectations are mildly increasing over the coming years. Forward rates out to 10 year or so increase much more rapidly than expectations, however, due to the risk premium of 12.5 basis points per year. For longer terms, however, the (negative) convexity term grows rapidly, not only moderating the impacts of expectations and convexity but also actually causing forward rates to decline with term.

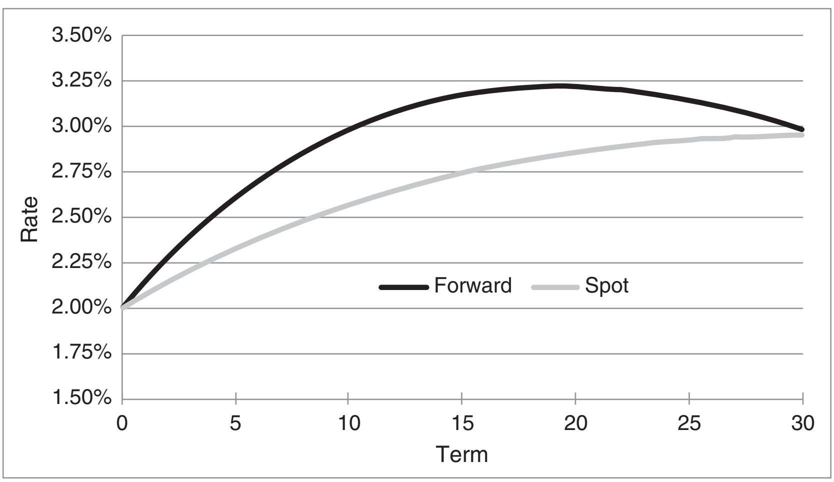

FIGURE 9.2 Continuously Compounded Forward and Spot Rates in the Vasicek Model. Model Parameters Are $r_0 = 2\%$, $\theta = 11\%$, $k = 0.0165$, and $\sigma = 0.95\%$.

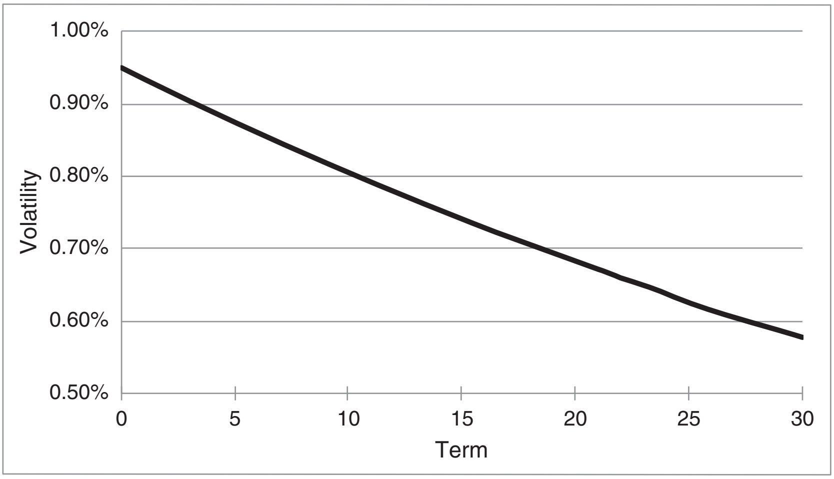

FIGURE 9.3 Term Structure of Forward Rate Volatilities in the Vasicek Model. Parameters Are $r_0 = 2\%$, $\theta = 11\%$, $k = 0.0165$, and $\sigma = 0.95\%$.

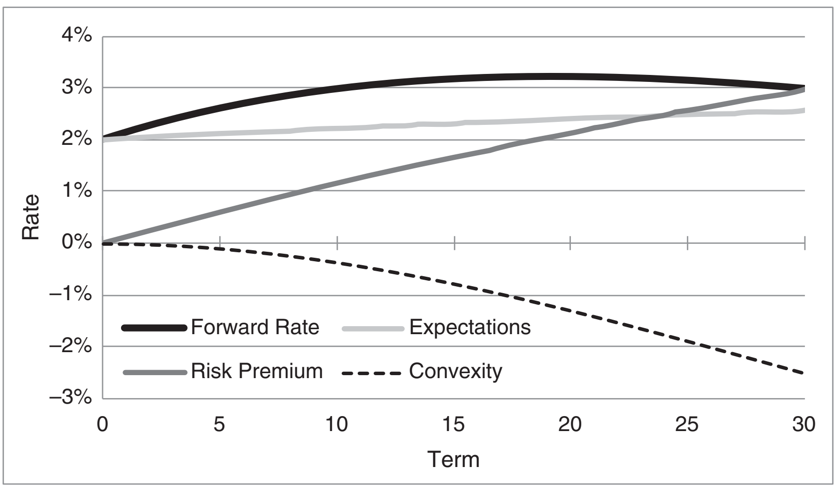

FIGURE 9.4 Decomposition of the Forward Rates in the Vasicek Model into Expectations, Risk Premium, and Convexity. Model Parameters Are $r_0 = 2\%$, $\lambda = 0.125\%$, $r_\infty = 3.424\%$, $k = 0.0165$, and $\sigma = 0.95\%$.

The Vasicek model has some limited uses for practitioners. It is a relatively simple model, which is a great advantage. Furthermore, as explained in Chapter 6, a single factor can explain a large fraction of term structure variability, particular across longer maturities. The parameters r_0, k, and θ can be jointly calibrated to approximate both the shape of the term structure and the shape of rate sensitivities to the factor (i.e., Figure 9.3). The parameter σ can be used to approximate an implied option volatility at one point of the term structure. With these considerations in mind, the model can reasonably be used, for example, to price, compare values, and hedge long-term bonds that are first callable after some intermediate number of years. The model is flexible enough to match the prices of noncallable bonds from 10 to 30 years, and also to match the most relevant volatility, namely, the volatility of 10-year rates. Bond sensitivities to changes in interest rates, defined in the model as changes in r, can be computed by shifting r_0, recomputing prices, and computing DV01s or durations. To the extent that bonds of one maturity are hedged with bonds of another maturity, the effectiveness of the resulting hedges depend on the reliability of the shape in Figure 9.3.

The model might also be used for trading and hedging relatively long-term bonds. The value of r_0 might be set each day so that the model 10-year rate matches the market 10-year rate. Deviations of calculated prices of longer-term bonds from model predictions might then be taken as signals of relative value, with hedge ratios calculated as described in the previous paragraph. The success of relative value trading along these lines depends on the extent to which the model captures the equilibrium or steady state of the shape of the term structure. Also, as before, hedging effectiveness depends on the reliability of the term structure of sensitivities to the factor.

The Vasicek model is not very widely used, however, because it is too simple for most applications. First, the model is not flexible enough to approximate the wide variety of observed market term structures. Or, put another way, calibrating the model to match one point on the term structure relies too heavily on the model to approximate all other points on the term structure. Second, while one factor does explain a lot of term structure variability, a second factor can capture significantly more variability. In terms of the discussion in Chapter 6, more than one factor is needed to hedge intermediate- and shorter-term bonds. Third, the Vasicek model cannot capture the empirical regularity, mentioned already, that the term structure of volatilities and, therefore, the term structure of factor sensitivities, typically rises quickly with term and then flattens or declines. This limitation means that the model cannot simultaneously handle options or other volatility-sensitive products that are spread out across the term structure, nor can it reliably hedge bonds most sensitive to one segment of the term structure with bonds most sensitive to another segment. With these caveats, the text turns to a much more flexible term structure model.

9.2 THE GAUSS+ MODEL

The Gauss+ model is well-known among practitioners as a tool for proprietary trading and hedging. The assumptions of the model are intuitively appealing, and they reasonably balance the goals of tractability and of capturing the empirical complexity of term structure dynamics. The goal of the text is to introduce the model, both in theory – through its equations and solution – and in application – through a full estimation of its parameters using recent data. Introducing a detailed estimation procedure here is noteworthy, because methodologies across the industry vary widely. The goal of the extensive appendix to this section is to enable a determined reader to estimate and implement the model independently.

The dynamics of the *cascade form* of the model are given in Equations (9.9) through (9.12). The factors r, m, and l denote the short-term rate of interest, a medium-term factor, and a long-term factor, respectively. The parameters μ and ρ are discussed presently. The mean reversion parameters of the factors are, α_r, α_m, and α_l, respectively, and the volatility parameters for the medium- and long-term factors are σ_m and σ_l, respectively. The two random variables in the model are dW^1 and dW^2. The subscript t denotes time-t observations of the factors, of changes in the factors, and of the random variables. Finally, then, the equations are,

$$dr_t = -\alpha_r(m_t - r_t)dt \tag{9.9}$$

$$dm_t = -\alpha_m(l_t - m_t)dt + \sigma_m(\rho dW_t^1 + \sqrt{1-\rho^2}dW_t^2) \tag{9.10}$$

$$dl_t = -\alpha_l(\mu - l_t)dt + \sigma_l dW_t^1 \tag{9.11}$$

$$E[dW_t^1 dW_t^2] = 0 \tag{9.12}$$

Given the structure of the model, it turns out that the medium- and long-term factors can be thought of as rates. The short-term rate mean reverts to the medium-term factor, which is meant to reflect business cycles and monetary policy factors. The medium-term factor reverts to the long-term factor, which is meant to reflect long-term expectations of inflation and the real interest rate, which ultimately depend on long-term trends in demographics, production technology, and so forth. And the long-term factor reverts to a constant, μ, which, as in the Vasicek model, can be thought of as including both a long-term expectation of the short-term rate and a risk premium. The mean reversion parameters are expected to be consistent with these economic interpretations; that is, the short-term rate reverts quickly to the medium-term factor; the medium-term factor reverts more slowly to the long-term factor; and the long-term factor reverts slowest of all to its target.

While the medium- and long-term factors trend as described in the previous paragraph, they also fluctuate around these trends. With respect to the evolution of the long-term factor in Equation (9.11), the fluctuation over a short time dt is $\sigma_l dW_t^1$. Because dW_t^1 is normally distributed with mean zero and standard deviation $\sqrt{dt}$, the instantaneous fluctuation of the long-term factor around its trend has mean zero and volatility $\sigma_l\sqrt{dt}$. The random terms in Equation (9.10) look complicated, but they simply ensure that the instantaneous fluctuation of the medium-term factor around its trend has a volatility of $\sigma_m\sqrt{dt}$ and a correlation of ρ with the fluctuation of the long-term factor around its trend. To see this, note that dW_t^2 also has mean zero and standard deviation $\sqrt{dt}$, and, from Equation (9.12), zero correlation with dW_t^1. It then follows from Equation (9.10) that the standard deviation of dm_t is,

$$\sqrt{\sigma_m^2(\rho^2 dt + [1-\rho^2]dt)} = \sigma_m\sqrt{dt} \tag{9.13}$$

that the covariance of dm and dl is,

$$Cov[\sigma_m(\rho dW^1 + \sqrt{1-\rho^2}dW^2), \sigma_l dW^1] = \rho\sigma_m\sigma_l dt \tag{9.14}$$

and, therefore, that the correlation of dm and dl is,

$$\frac{\rho\sigma_m\sigma_l dt}{\sigma_m\sqrt{dt}\times\sigma_l\sqrt{dt}} = \rho \tag{9.15}$$

The evolution of the short-term rate in the model, Equation (9.9), is meant to reflect how central banks conduct rate policy. The Fed, for example, keeps the short-term policy rate pegged or fixed at a target, but moves that target over time in a manner deemed appropriate for the state of the business cycle and monetary conditions. Mathematically, in Equation (9.9), the short-term rate is fixed over the very short time interval, dt, in the sense that there is no random variable shocking the dynamics of r. The rate, r, is pushed gradually, however, toward the medium-term factor, m, which in turn reverts to the long factor, l.

The medium- and long-term factors move expectations of the short-term rate as of future dates. Because these expectations move smoothly over time, the medium- and long-term factors are assumed to move in a continuous fashion. The short-term rate, by contrast, changes in the real world by discrete amounts, on a set of fixed dates, according to central bank policy decisions. The model approximates the future outcomes of this process, however, by a continuous process starting at today's short-term rate.

The lack of a volatility term in Equation (9.9) is an important feature of the Gauss+ model. As pointed out in the discussion of the Vasicek model, mean-reverting factors generate a downward-sloping term structure of volatility. Largely because of central banks, however, empirical and implied term structures of volatility tend to have a hump; that is, volatility is low for very short-term rates before increasing to a peak at intermediate-term or longer rates. In the Gauss+ model, the lack of a random shock in the dynamics of r keeps short-term rate volatility low. In this way, the Gauss+ model can match empirically observed hump-shaped term structures of volatility.

In passing, the Gauss+ model gets its name from the lack of a volatility term in Equation (9.9). The "Gauss" part of the name indicates that interest rates have a normal or Gaussian distribution. But while most one-, two-, or three-factor normal models have a corresponding number of sources of risk, the Gauss+ model, strictly speaking, has three factors, but only two sources of risk. The model has three *state variables* in the sense that describing the state of the world in the model requires knowing the three factors, r, m, and l. There are only two sources of risk, however, namely, dW^1 and dW^2. The "+" in the name, therefore, indicates the somewhat unusual presence of a factor or state variable that is not also a source of risk.

As a final comment on the structure of the model, Equations (9.9) through (9.12) are the risk-neutral dynamics of the model. Additional assumptions allow for the identification of an implicit risk premium. An application of the model developed next shows how this may be done, assuming that only the long-term factor earns a risk premium.[2]

Appendix A9.2 gives a detailed account of how the parameters of the Gauss+ model might be estimated from bond or swap data. A simplified overview is presented here, using daily data from January 2014 to January 2022 on the fed funds target rate (see Chapter 12) and on zero coupon bond prices of various maturities, which are derived from the prices of US Treasury bonds. The time series of zero coupon bond prices are published by the Federal Reserve Bank of New York and are publicly available.[3]

Consistent with the interpretation of the model, the short-term rate, r, is taken each day as equal to the fed funds target rate on that day. While a general collateral repo rate is theoretically more consistent with a term structure of Treasury interest rates, repo rates exhibit occasional idiosyncratic jumps that complicate the estimation without significant offsetting advantages.

Once estimated, the model factors are chosen to "fit" or match, each day, the one-year rates, two and 10 years forward, and the fed funds target rate. Furthermore, as discussed presently, with these two- and 10-year forward rates fair by construction, the model becomes a tool for trading value in other parts of the curve relative to these fitted points. For ease of exposition, by the way, all forward rates mentioned from this point denote one-year rates some number of years forward.

---

[2]See, for example, Cochrane, J., and Piazzesi, M. (2009), "Decomposing the Yield Curve," *AFA 2010 Atlanta Meetings Paper*, January 26.

[3]See Gürkaynak, R., Sack, B., and Wright, J. (2006), "The US Treasury Yield Curve: 1961 to the Present." Data are available at https://www.federalreserve.gov/data/nominal-yield-curve.html.

Although the model ultimately fits the two- and 10-year forward rates, note that the factors m and l are in no way those forward rates themselves, in the way that r equals the fed funds rate. The relationships of all forward rates to m and l depend on the estimated model parameters and, in particular, on the mean reversion parameters. More specifically, the larger α_r, the faster changes to the factors m and l make their way into the term structure of rates. The larger α_m, the faster m converges to l, and the less m affects longer-term yields. And the smaller α_l, the slower l converges to μ, and the more similar or parallel is the effect of l on all longer-term yields.

9.3 A PRACTICAL ESTIMATION METHOD

The estimation method presented here proceeds in stages, with each stage estimating a subset of the model's parameters.[4]

Figure 9.5 shows the coefficients of regressing the changes in the zero yields of various terms on changes in the two-year zero yield – the dark gray bars – and on changes in the 10-year zero yield – the light gray bars. For example, the regression of the three-year yield gives a coefficient of 0.91 on the two-year yield and 0.22 on the 10-year yield. By contrast, the regression of the 15-year yield gives a coefficient of −0.20 on the two-year yield and 1.10 on the 10-year yield. In any case, as explained before, all these regression coefficients implicitly describe the mean reversion parameters of the model. This stage of the estimation, therefore, chooses the parameters α_r, α_m, and α_l so that the model captures the empirically observed regression coefficients as closely as possible. This staging is possible, because, as shown in the appendix, the model regression coefficients depend only on the mean reversion parameters, not on the volatility parameters. In any case, the mottled dark and light gray bars in Figure 9.5 show the success of this stage of the estimation. There are a set of mean reversion parameters, given hereafter, such that implied regression coefficients from the model very closely match the empirically observed regression coefficients. And, while not reported here, these matches are within statistical confidence intervals.

The next stage of the estimation is to find the volatility and correlation parameters, σ_m, σ_l, and ρ, so that the term structure of volatilities in the model matches the term structure of volatilities in the data as closely as possible. The result of this optimization is shown in Figure 9.6. Once again, the model is flexible enough to do an excellent job of matching empirical properties of the term structure of interest rates.

[4]This approach is significantly easier to implement than the maximum likelihood methods that are standard in the term structure literature.

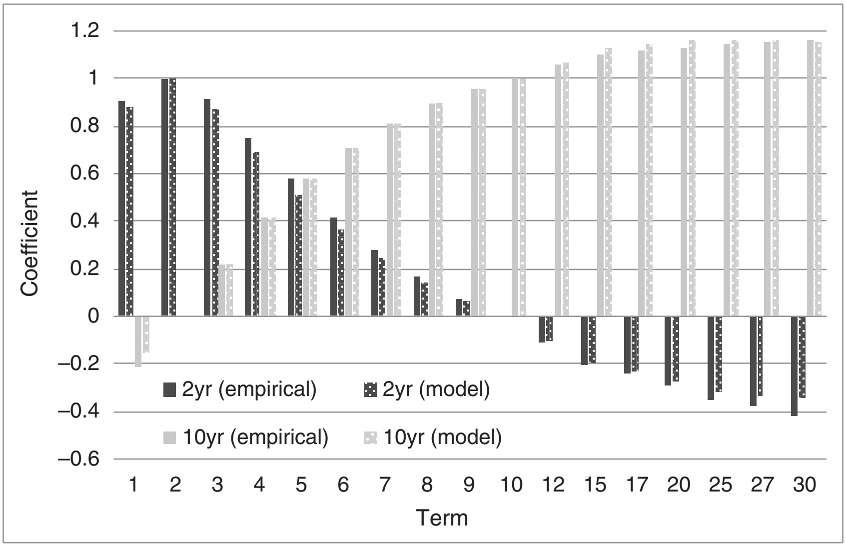

FIGURE 9.5 Coefficients of Regressing Zero Coupon Bond Yields of Various Terms on Two- and 10-Year Zero Coupon Bond Yields, from Empirical Analysis and as Implied by the Estimated Gauss+ Model.

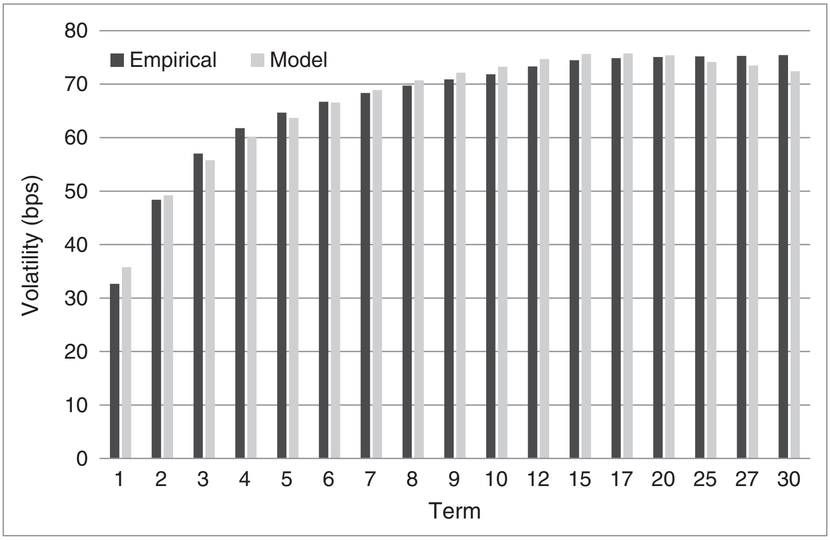

FIGURE 9.6 Yield Volatility in Annual Basis Points, from Empirical Analysis and as Implied by the Estimated Gauss+ Model.

The last remaining parameter to be estimated is μ, the value to which the short-term rate reverts, over the very long-term. The estimation procedure suggested here finds the μ that minimizes the sum of the squared errors of observed yields relative to model yields across the whole data sample.

Following the estimation procedure described, Table 9.1 reports the resulting Gauss+ parameter values. The mean reversion parameters are in the order expected, with the central bank reaction fastest, the speed of the medium-term factor's reversion to the long-term factor next, and the speed of the long-term factor's reversion to μ slowest. Alternatively, the time for each process to converge halfway to its target, given in the half-life column, is about eight months for r, 13 months for m, and 42 years for l. Because of the mean reverting nature of the factors, the volatility parameters of 109 and 96 basis points for the medium- and long-term factors, respectively, translate into the lower zero yield volatilities shown in Figure 9.6. The parameter μ, as the very long-run target for the short-term rate, might seem high at over 10%, but the long-term factor reverts very slowly to this target, and that target includes a risk premium, which is discussed further next.

The term-structure properties of the estimated model are well described by Figure 9.7, which graphs the change in forward rates for a change in each of the factors as a function of term. For example, the seven-year forward, changes by 0.9 basis points for every basis point change in the long-term factor. Taken as a whole, the figure shows that short-term factor affects only the very short end of the curve. The medium-term factor, which drives the two- to three-year part of the curve, can be thought of as capturing monetary policy in the sense of encapsulating where the market believes the short-term rate will be in two to three years. The long-term factor has its biggest impact in six to eight years and, beyond 10 years, is the sole factor driving forward rate changes and volatilities.

The time series properties of the model can be described by graphing its factors over time. As mentioned already, the short-term rate is set each day to

TABLE 9.1 Estimated Parameters of the Gauss+ Model from US Treasury Zero Coupon Yields, January 2014 to January 2022. Half-Life Is in Years.

Parameter	Estimate	Half-Life
α_r	1.0547	0.66
α_m	0.6358	1.09
α_l	0.0165	42.01
σ_m	109.2 bps	
σ_l	96.4 bps	
ρ	0.212	
μ	10.555%	

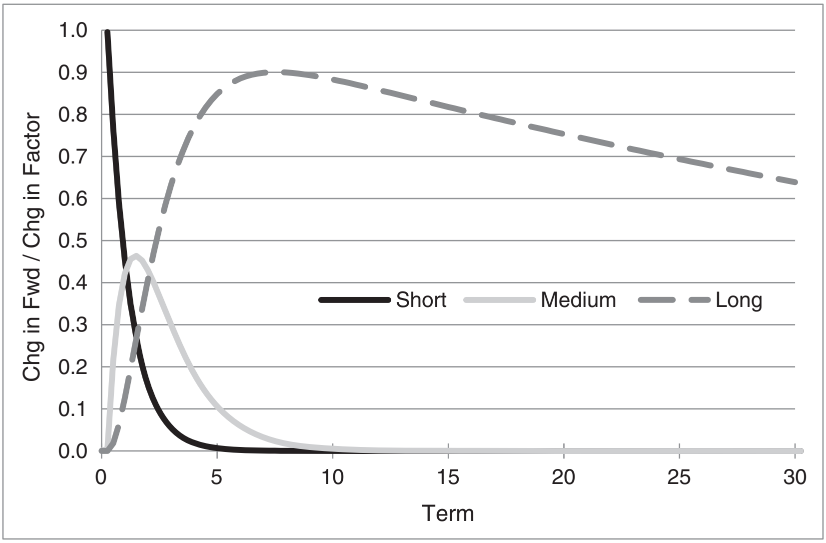

FIGURE 9.7 Changes in Forward Rates Relative to Changes in the Short-Rate, the Medium-Term Factor, and the Long-Term Factor in the Estimated Gauss+ Model.

the fed funds target rate, and the medium- and long-term factors are set so as to match the model and market two- and 10-year forward rates. Figure 9.8 graphs these empirically recovered market factors from January 2007 to January 2022.[5] The two-year forward rate is included in the graph to focus the interpretation of the medium-term factor. When rates are high, this factor closely tracks the two-year forward rate, confirming that the medium-term factor loosely corresponds to where the market expects the short-term rate to be in two years. When rates are low, however, near the zero lower bound, the medium-term factor can fall into deeply negative territory. In this sense, the medium-term factor is a "forward shadow rate" that reflects future expectations of the short-term rate and that can be traded explicitly in the Gauss+ model. This interpretation differs from models in which the "shadow rate" is what the short-term rate would be now if it were not bounded above zero.[6]

[5]The model is estimated using data from January 2014, but the resulting parameters are used to extract model factors back through 2007. Also, instead of the long-term factor itself, the figure graphs the long-term factor shifted forward 10 years. This allows the series to be more easily interpreted as an approximation for expectations of the short-term rate in 10 years. See Appendix A9.2 for further details.

[6]See, for example, Wu, J., and Xia, F. (2016), "Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound," *Journal of Money Credit and Banking* 48, March-April; Bauer, M., and Rudebusch, G. (2016), "Monetary Policy Expectations at the Lower Bound," *Journal of Money Credit and Banking*

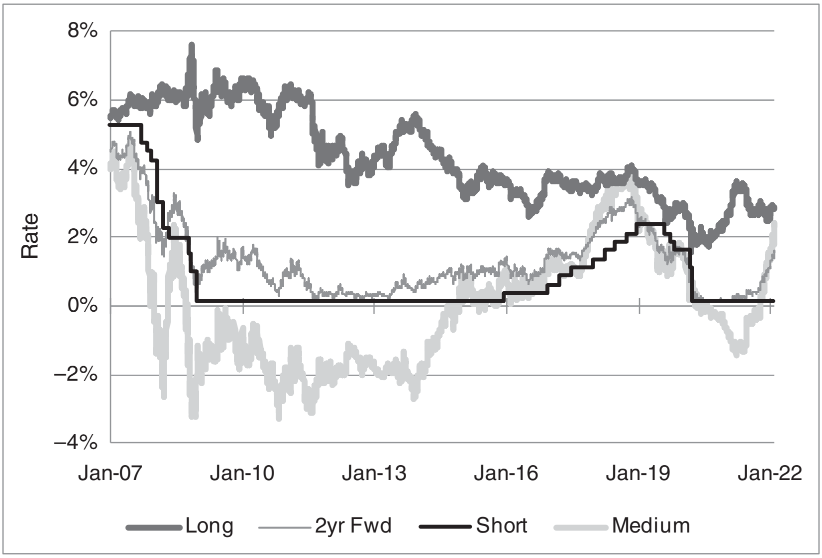

FIGURE 9.8 The Two-Year Forward Rate and Gauss+ Factors Extracted from Daily Market Data.

In any case, the nature of the medium-rate factor as a leading indicator of the short-term rate can be seen by comparing those two series in Figure 9.8. Particularly striking is the steep increase in the medium-term factor starting in early 2014, a couple of years before the Federal Reserve began raising rates.

9.4 RELATIVE VALUE AND MACRO-STYLE TRADING WITH THE GAUSS+ MODEL

In the context of a term structure model, a relative value trade is one that is not exposed to or hedged against changes in the factors. Ideally, a trader would find an individual security that is cheap relative to the model and, from various analyses, is expected to revert soon to being fair to the model. The trader would then buy that security; hedge some or all of its factor exposure by selling fair or rich securities in the same part of the curve; and earn the resulting profits. In practice, however, individual securities are often persistently cheap or rich to the model, so that convergence is not expected

48, October; and Kim, D., and Singleton, K. (2012), "Term Structure Models and the Zero Bound: An Investigation of Japanese Yields," *Journal of Econometrics* 170, September.

in the short run, and securities neighboring in term are often all fair or rich together. Most of the time, therefore, relative value opportunities arise in which a trader receives forward rates in one part of the curve that are high relative to the model but expected to converge promptly to the model; pays forward rates in another part of the curve that are too low, but expected to converge promptly; and structures trades to minimize factor exposures. In this synopsis, the speed at which any detected mispricing is likely to correct itself is an important trading consideration.

As an example of using the Gauss+ model along these lines, consider the following framework. Compute the time series of the nine-year forward rate minus the model-predicted equivalent rate, where each observation can be called a fitting error. Then construct a signal from this times series as the difference between its five-day and 40-day moving averages. If the demeaned fitting error at a given time is positive, so that the nine-year forward is too high relative to the model, and if the signal is negative, so that fitting errors have started to fall, that is, they have started to converge to the model, then receiving at that forward rate might be considered an attractive trade.

The problem with receiving or paying the nine-year forward in isolation, however, is that it has significant exposure to the medium- and long-term factors, which is not consistent with the spirit of relative value trading. One solution is to pool together and size several attractive relative value trades so that their exposure to the factors is minimal. In addition, traders can diversify across mean reverting trades. It turns out, for example, that nine- and five-year fitting errors tend to be positively correlated. This fact makes it attractive to pay in one rate and receive in the other. Figure 9.9 shows, in fact, that the difference between the nine- and five-year signals is strongly mean reverting, which is one of the most important properties of relative value trades.

Unlike relative value trading, which finds value in the absence of factor exposure, macro trading takes direct or indirect views on the factors. Simple examples include positions based on predictions of changes in rates or in the slope of the term structure that differ from what is priced in the market. A more complex example, which has attracted more interest over time, is trying to trade the long-run level of the short-term rate. As explained earlier, long-term forward rates are a combination of expectations, risk premium, and convexity. Within the structure of the Gauss+ model, with the help of a strategic assumption, long-term forward rates can be separated into these three components. A macro trader can then decide that long-term forwards are too low or too high and position accordingly. The details of this decomposition of forward rates are complex and, therefore, appear in the appendix to this chapter. The text continues with an intuitive approach.

In the context of any term structure model, it is relatively straightforward to determine the effect of convexity on forward rates. It is much more

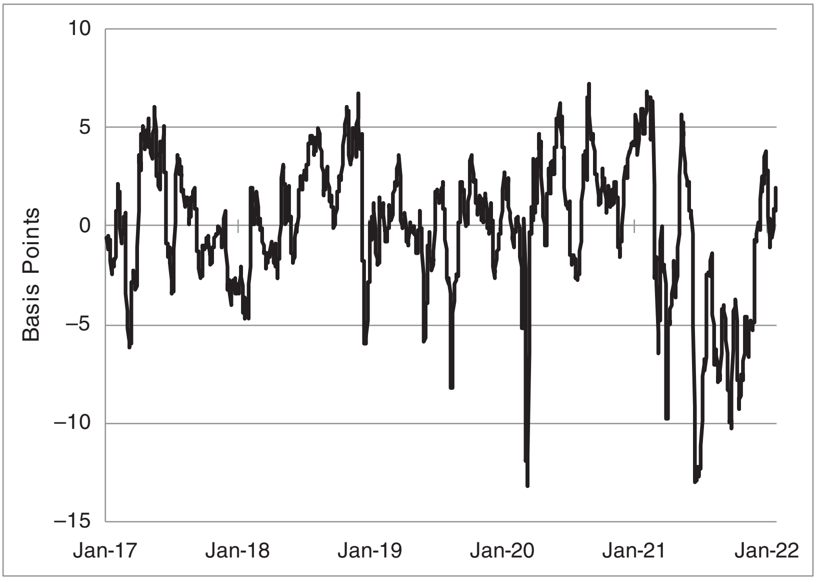

FIGURE 9.9 Difference Between the Nine- and Five-Year Signals. Each Signal Is Based on a Difference of Moving Averages of Deviations of Market from Gauss+ Model Rates.

difficult, however, to separate expectations from risk premium. A number of approaches appear in the academic literature, but these are not without various shortcomings.[7] The method proposed here relies on one key assumption: expectations of future short-term rates do not change past some point in the future. Anyone with a view of the short-term rate in 15 years, for example, has the same view of the short-term rate in 20 or in 30 years. Consequently, any difference in forward rates beyond some term is attributable not to rate expectations, but solely to risk premium and convexity. In this way, expectations and risk premium can be separated and calculated from observable rates.

Using the assumption just described and the parameters of the Gauss+ model estimated previously, Figure 9.10 graphs the risk premium on the

---

[7]See, for example, Adrian, T., Crump, R., and Moench, E. (2013), "Pricing the Term Structure with Linear Regressions," *Journal of Financial Economics* 110(1), October; Ang, A., and Piazzesi, M. (2003), "A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables," *Journal of Monetary Economics* 50(4); Cieslak, A. (2018), "Short-Rate Expectations and Unexpected Returns in Treasury Bonds," *Review of Financial Studies* 31(9); and Kim, D., and Orphanides, A. (2012), "Term Structure Estimation with Survey Data on Interest Rate Forecasts," *Journal of Financial and Quantitative Analysis* 47(1).

10-year forward rate over time, measured along the left axis. A value of 60 basis points on any day, for example, means that, on that day, 60 basis points of the 10-year forward rate is attributable to risk premium. The lighter line is the level of the 10-year forward rate, measured along the right axis. The risk premium often moves in the opposite direction of rates. As rates decline, the term structure typically steepens, which the model interprets as an increase in risk premium. This observed behavior is consistent with the low-inflation regime over the last several decades, in which government bond prices increase as other risky assets fall in value. In that environment, as rates fall, and have less room to fall even further, bonds are less able to hedge the declining value of other assets and, as a result, are worth less themselves.

The flip-side of identifying the risk premium, of course, is identifying the long-run expectation of the short-term rate in the estimated Gauss+ model, more specifically, the 10-year forward rate minus the term-appropriate risk premium plus the term-appropriate convexity. The time series of this expectation is shown in Figure 9.11, along with a different estimate, formed from real rate forecasts and inflation estimates at the Federal Reserve Bank of Cleveland.[8] While the model and outside series track each other quite well over time, there are trading opportunities to use the difference between the two series as a measure of value. Put another way, the difference between the

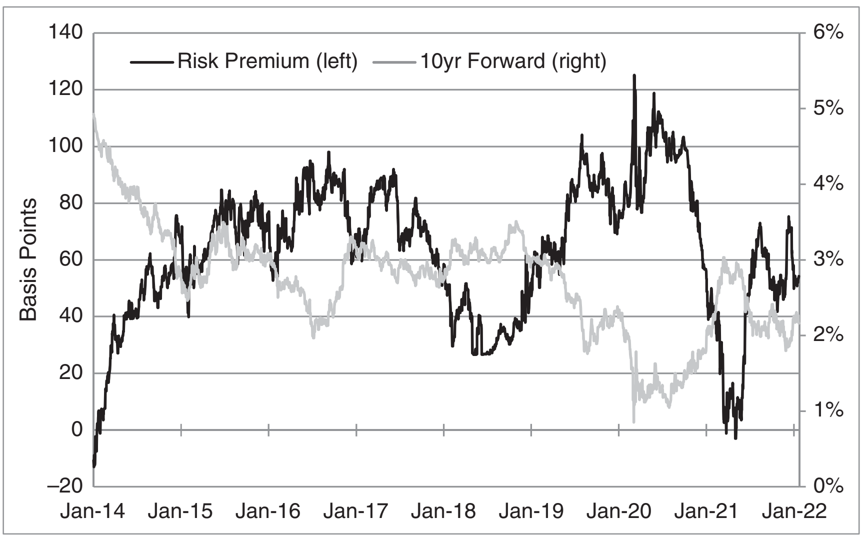

FIGURE 9.10 Estimated Risk Premium on the 10-Year Forward Rate.

[8]The construction of the Cleveland Fed series is described further in the appendix to this chapter.

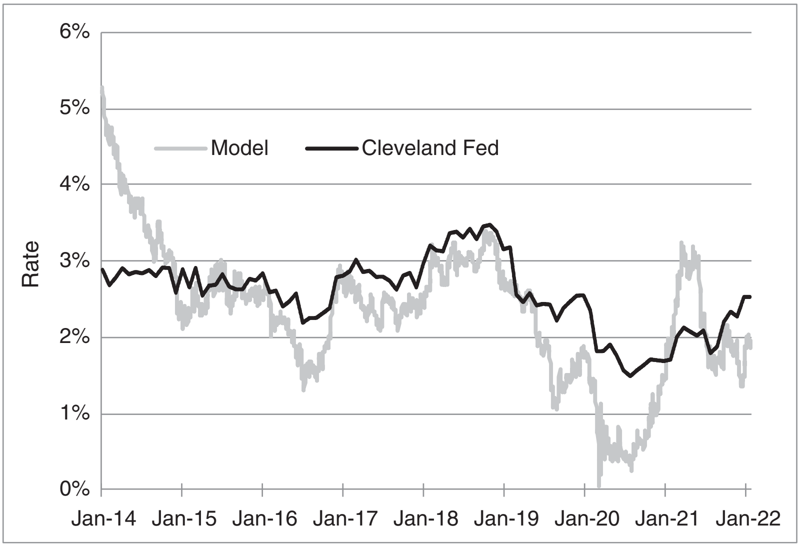

FIGURE 9.11 Long-Run Expectations of the Short-Term Rate, as Implied by Gauss+ Fitted to Market Rates and by Fundamental Analysis at the Federal Reserve Bank of Cleveland.

Gauss+ market-implied view – the long-run rate priced in the market – and the exogenous, economist-generated, fundamental view – what one thinks the long-run rate should be – can be used as a basis for taking outright long or short positions in bonds.

CHAPTER 10

Repurchase Agreements and Financing

The *repurchase agreement* or *repo* market might be the most important market about which most people know little or nothing. Money market funds (MMFs) and other investors rely on repo as a short-term, liquid asset; broker–dealers and other financial entities use repo to fund their inventory of securities; repo enables market participants to take short positions in fixed income markets; and, as described in Chapter O, the Federal Reserve has historically and continues presently to use repo to conduct monetary policy. Finally, and most recently, interest rate derivatives and many loan products are transitioning from the *London Interbank Offered Rate (LIBOR)* indexes to the *Secured Overnight Financing Rate (SOFR)*, which is derived from rates on repo transactions.

The first few sections of the chapter describe repurchase agreements, the uses of repo, and the structure and size of the market. A short section then describes the computation of the recently prominent SOFR, a rate featured in Chapters 2, 12, and 13. The subsequent section explains some of the determinants of the interest rates on both *general collateral (GC)* and *special* repo transactions. The penultimate section discusses repo in the context of financing risk, along with relevant changes in banking regulation. The final section is a case study of how MF Global fell in large part due to its inappropriately sized "repo-to-maturity" trades.

10.1 REPURCHASE AGREEMENTS

The most straightforward description of a *repurchase agreement* or *repo* is as a secured loan with bonds or other financial instruments as collateral. Figures 10.1 and 10.2 depict an example in which one counterparty – the "repo seller" – borrows \$109,898,438 for 44 days at a rate of 0.015%, while giving \$100 million of the 2.75s of 02/15/2028 as collateral. At the time of the trade, the price of the bond is 109-28 3/4, that is, $109 + 28.75/32 = 109.898438$ per 100 face amount, which means that the value of the collateral at the start of the trade equals the amount loaned. In practice, the

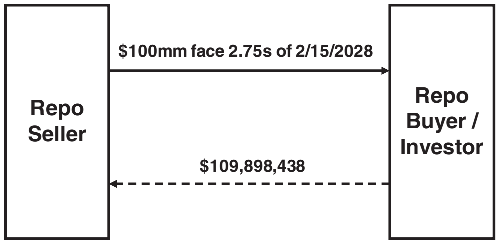

FIGURE 10.1 Initiation of a Repurchase Agreement.

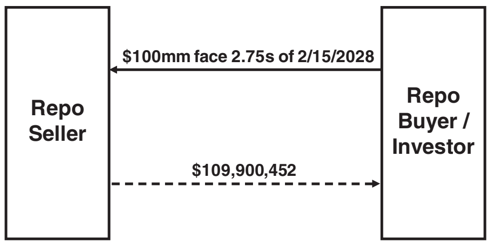

FIGURE 10.2 Unwind of a Repurchase Agreement.

value of the collateral typically exceeds the amount loaned, and this feature of repo is discussed next. Continuing with the present example, however, Figure 10.1 shows the flows of cash and bonds at the initiation of the trade, and Figure 10.2 shows the flows at the expiration or unwind of the trade. In the latter, the repo seller pays principal plus interest, $\$109{,}898{,}438 \times (1 + 0.015\% \times 44/360)$, or $109,900,452, to discharge the loan, and then takes back its $100 million face amount of bonds.

The lender of cash in Figures 10.1 and 10.2 is protected by the bonds taken as collateral against the loans: if the borrower of cash defaults, the lender of cash has a claim on those bonds. But lenders of cash in the repo market are protected even further by a *safe harbor* from the bankruptcy code, by a *haircut* or *initial margin*, and by *variation margin*.

With respect to the safe harbor, if the borrower of cash defaults on its loan, the lender may immediately liquidate the bonds held as collateral to recover the loan amount, returning any realized excess cash to the borrower. This is very different from the general treatment of secured loans, in which lenders are *stayed* or prevented from liquidating the collateral of defaulting borrowers until given permission to do so by a bankruptcy court. It is for this reason, in fact, that repo trades are legally structured as repurchase agreements and not as secured loans. Legally, the borrower of cash, that is, the repo seller, sells the bonds to the repo buyer at the start of the trade and agrees, at unwind, to repurchase the bonds at a slightly higher, prespecified price. While economically the same as a secured loan, this legal structure casts the unwind as the settlement of a securities trade, which enjoys a safe harbor from bankruptcy rules.[1]

In addition to the safe harbor, lenders of cash in the repo market are typically protected by *haircuts* or *initial margin*. At a haircut of 3% for example, collateral value is reduced by 3% when considered against the loan amount. Therefore, a repo seller can borrow only $97 against a bond worth $100, leaving a margin of $3, that is, an excess of collateral value over loan value of $3. In the example of this section, with a haircut of 3%, the cash borrower raises only 97% of $109,898,438, or $106,601,485, against the $100 million of the 2.75s of 02/15/2028, leaving initial margin of $3,296,953. With this buffer, if the borrower of cash defaults and the bonds fall in value from $109.9 million to, say, $108 million, the lender still has more than enough collateral value to cover the outstanding loan amount of $106.6 million.

Variation margin, the last of the protections mentioned earlier, requires borrowers to maintain collateral value as prices change so as to keep margin constant. Continuing with the example in the previous paragraph, if the bonds' value falls from $109.9 million to $108 million against the $106.6 million loan, the cash borrower receives a *margin call* to post additional bonds with a value of $1.9 million to restore margin to the original $3.3 million. Alternatively, the borrower can satisfy the call by paying the lender $1.8 million in cash, which reduces the loan amount from $106.6 million to $104.8 million, which is appropriately collateralized by $108 million of bond value at a haircut of 3%. If, on the other hand, bond value increases, the cash borrower can take back posted collateral or cash in excess of margin requirements. Variation margin calls in the repo market are typically issued daily.

Summarizing the protections in a repo agreement, a lender of cash suffers a loss only if the borrower defaults at the same time that collateral declines in value by more than posted margin.[2]

---

[1]To understand the logic of the general bankruptcy stay, consider the case in which a manufacturer of goods has borrowed money on the collateral of machines in its plant. If, in the event of a default, secured creditors could immediately seize those machines, all other creditors and equity holders would likely be irreparably harmed, whereas a limited delay could easily result in an orderly reorganization that maximizes the ongoing or liquidation value of the manufacturer. In contrast, it has been argued that financial markets are best served by the safe harbor, which minimizes the uncertainty of settlement.

[2]With government bond collateral, it can be argued that this exposure is *right-way risk*: government bonds tend to do well during the market upheavals that typically characterize counterparty defaults. On the other hand, with corporate or certain mortgage collateral, it can be argued that repo exposure to default is *wrong-way risk* because counterparty defaults tend to happen in exactly those scenarios in which collateral values fall.

10.2 USES OF REPURCHASE AGREEMENTS

This section describes the four major uses of repurchase agreements: investing and cash management, long financing, shorts or short financing, and collateral swaps.

Repo is popular as a short-term investment and as a place to park cash for a short period of time. Not only is lending cash through repo particularly safe, as explained in the previous section, but lenders can calibrate the default risks they are willing to bear by choosing haircuts (e.g., 2% for Treasuries, 5% for investment-grade corporates), the set of acceptable collateral (e.g., government bonds with maturities less than two years, any government bond, government bonds and investment-grade corporates), and their repo counterparties. Furthermore, by choosing the term of the repo agreement (e.g., overnight, 30 days), lenders can choose their desired liquidity, that is, the time after which they will again have use of their cash. Along these lines, repo is an important alternative to bank deposits. Both repo and bank deposits are highly liquid, but, while bank deposits of up to $250,000 are federally insured, larger deposits at many banks might not be considered as safe as secured repo investments.

Because of repo's popularity as a short-term investment, many repo buyers are called *repo investors*. The most important class of such investors are money market funds, which are described in detail in Chapter O. Because shareholders in these funds can easily redeem their shares for cash except, perhaps, in times of great stress, managers of money market funds maintain liquidity by investing a large portion of their assets in particularly liquid products, like repo. Mutual funds, which must also plan for some amount of shareholder redemptions, also invest in repo as a liquid investment. Other repo investors, like municipalities and nonfinancial corporations, are motivated by the need to manage cash inflows and outflows. A municipality receives tax payments and makes expenditures throughout the year, but the timing of receipts generally does not match the timing of outlays. Therefore, cash on hand needs to be invested in short-term, safe products, like repo. Similarly, because the receivables and payables of nonfinancial corporations do not generally coincide, repo is often used to park excess cash on hand.

A second common use of repo is *long financing*, which refers to the purchase of financial instruments with mostly borrowed money. An important example, illustrated in Figure 10.3, is the use of repo by broker–dealers to facilitate market making in bonds. As shown in the top panel of the figure, a customer sells bonds to the trading desk, for which the desk pays cash. But the business model of trading desks is to earn relatively small fees relatively frequently using a minimum of scarce capital. Therefore, rather than using capital to purchase the customer's bonds, the trading desk chooses to raise funds in the repo market, that is, to *repo out* the bonds, or sell the repo, to a repo buyer. Note that, because of haircuts, the trading desk cannot raise

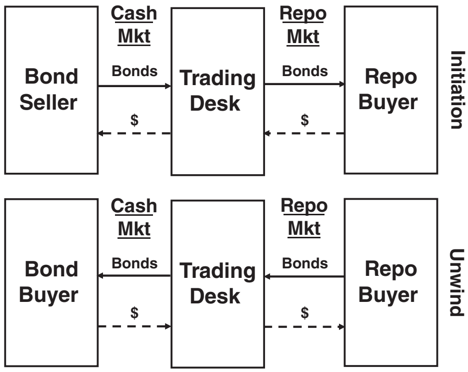

FIGURE 10.3 Long Financing.

the full value of the bonds it bought in the repo market. Consequently, the trading desk does use a small amount of its own capital to fund margin, that is, to pay the customer that portion of the purchase price that is not funded with repo. In short, a fixed income market-making desk funds its inventory predominantly with repo, but in small part with capital.

The bottom panel of Figure 10.3 shows how the trading desk exits its cash and repo market positions. At some point after the desk's original purchase of the bonds, another customer comes along wanting to buy those same bonds. The trading desk sells the bonds to the customer, uses the proceeds to pay off its repo borrowings and takes back the bonds for delivery to the customer. At this point it can be pointed out that purchases and sales of a particular bond on a given day cancel, and do not require repo funding. It is only the residual inventory carried from one day to the next that needs to be funded.[3]

Another example of long financing would be a trader or investor who wants to lever returns by buying a bond with borrowed money. At initiation, the investor buys the bond in the cash market and repos it out to raise most of the purchase price. To terminate the trade, the investor sells the bond in the cash market and uses the proceeds to pay off the repo loan. Of course, if the bond price falls in the interim, the investor must raise cash elsewhere to discharge the loan. While the market-making desk described in the previous paragraph also seems to bear the risk of falling prices, market makers tend

[3]To be even more precise, the cash required for the settlement of trades on a given day is funded that day. For example, in the case of the Treasury market, which settles T+1, a bond purchased from a customer on Monday has to be paid for on Tuesday. Therefore, if that bond were not sold on Monday, the trading desk would need to raise funds on Tuesday morning to pay the customer later that day.

to hedge the interest rate risk of their inventory. Therefore, any loss on the long position, which leaves a deficit with respect to the repo loan, can likely be covered by the corresponding profits of a hedging, short position.

Sometimes the purchaser of a bond knows the exact term over which the bond needs to be financed. In most cases, however, the term of financing is not known. In the case of the market-making desk, the customer that ultimately purchases the bond from inventory may appear very soon after the desk itself purchases the bond or days later. In the case of the leveraged investor, market conditions might cause the investor to terminate the position sooner or later. Therefore, the repo agreement in long financing trades may expire before the end of the trade. In that case, the repo seller needs to *roll* the repo. Consider a repo seller who has financed the purchase of $100 million face amount of the 2.75s of 02/15/2028 with an overnight repo along the lines of Figure 10.1. If the repo seller wants to stay in the trade for an additional day, that is, if the repo seller does not yet want to repay the repo loan, the repo buyer would have to agree to roll the loan for another day. If the repo buyer does agree, collateral or cash is adjusted to maintain the appropriate amount of margin, and the repo seller can keep the money for another day. If the repo buyer does not agree, the repo seller must find another counterparty to finance the position and use the money borrowed from that counterparty to repay the loan due to the original repo buyer. The process of rolling a repo is depicted in Figure 10.4. Note that, because the price of the bonds may have changed over the day, the repo seller may receive more or less cash from the new repo buyer than is needed to discharge the loan due to the original repo buyer.

The duration of repo agreements can be overnight,[4] *term* (i.e., any fixed duration greater than one day), or *open*. Open repos are overnight repos

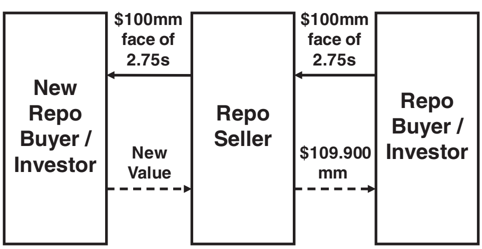

FIGURE 10.4 Roll of a Repurchase Agreement.

[4]Market participants say "overnight repo" rather than one-day repo, because, historically, cash was available to borrowers from the end of one day to the beginning of the next day. However, given reforms since the financial crisis of 2007–2009, these agreements are, in fact, be more accurately described as one-day repo.

that roll automatically for an additional day until cancelled by the lender of cash. The popularity of this contractual form reflects the many instances in which counterparties are unsure of how long financing will be needed.

At this point, it can be noted that operational aspects of repo markets are complex. Consider the roll in Figure 10.4. The repo seller needs to post its bonds to the new repo buyer to get cash, but the original repo buyer is unlikely to release those bonds until its loan is repaid – that is, after all, the whole point of holding collateral. Issues like this are resolved through the services of a third-party agent, as described in the next section.

A third common use of repo is the establishment of short positions or *short financing*. A short position is initiated by selling a bond without owning it, whether as a bet that its price will fall or as part of a larger trade to hedge a particular long position. In the top panel of Figure 10.5, a hedge fund sells bonds that it does not have to some buyer in the cash market. The hedge fund then takes the bonds it sold from a broker–dealer in the repo market, delivers those bonds to the buyer, receives the proceeds of the bond sale, and then lends those proceeds to the broker–dealer. To unwind its short position, as shown in the bottom panel of Figure 10.5, the hedge fund buys the bonds in the cash market, returns the bonds to the broker–dealer, collects its loan proceeds from the broker dealer, and uses those proceeds to pay for its purchase of the bonds in the cash market.

Using the language of this chapter, the hedge fund does buy the repo, but its motivation is very different from cash managers or investors who buy repo: the hedge fund's motivation is not to park or invest cash, but to borrow the bonds. For this reason, it is more commonly said that the hedge fund does a *reverse repo* or that it *reverses in* the bonds. Note that this industry jargon is such that a single transaction – the repo agreement

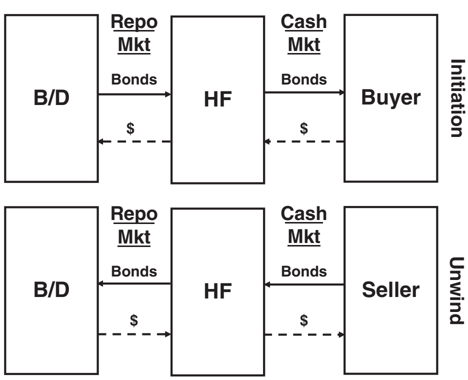

FIGURE 10.5 Short Financing.

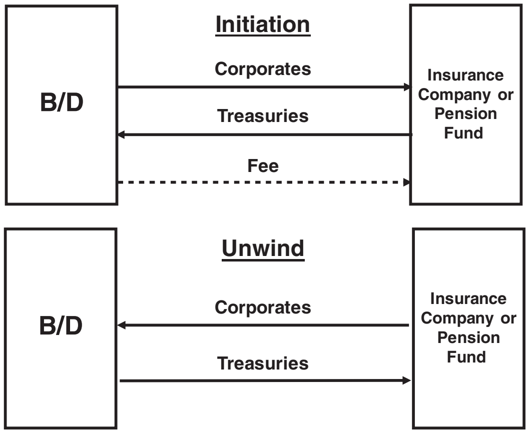

FIGURE 10.6 Collateral Swap.

between the broker–dealer and the hedge fund – changes names depending on the counterparty: the broker–dealer does a repo or repos out the bonds, while the hedge fund does a reverse repo or reverses in the bonds.

The different motivations for buying repo introduce another aspect of the repo market. As mentioned already, when cash managers and investors buy repo, they choose the general asset classes and maturities they are willing to accept as collateral. But they do not care about which particular bonds they receive. For this reason, these trades are called *general collateral* or *GC* trades, and the repo seller chooses which bonds to deliver subject to some set of general parameters. By contrast, a market participant who shorts a bond and does a reverse repo requires the exact bond that was sold in the cash market. These trades, therefore, are called *special collateral* or *specials* trades. The pricing of GC and special spreads is discussed later in the chapter.

The fourth and final use of repo discussed here is a *collateral swap*, which is described in Figure 10.6. A broker–dealer might choose to hold an inventory of corporate bonds but would like its overall holdings to be more liquid, that is, more easily and at a lower cost turned into cash, should the need arise. This preference for additional liquidity might result from prudent management, regulatory requirements, or both. In any case, at the same time, an insurance company or pension fund with long-term holdings of Treasury bonds might be happy to sacrifice liquidity temporarily in order to earn some extra return. The broker–dealer and the insurance or pension company, therefore, might agree on a collateral swap: the broker–dealer exchanges its corporates for Treasuries over the term of the trade, paying a fee to do so (e.g., 25 basis points on the value exchanged).

10.3 MARKET STRUCTURE AND SIZE

Table 10.1 summarizes the structure and size of the repo market, as of July 2021.[5] A main division of the market is between the *tri-party* and *bilateral* markets. In tri-party trades, an agent provides various collateral management services. In bilateral trades, the two counterparties to each repo transaction manage all of the operational details.

One of the key services of the tri-party system is to eliminate the risk that collateral is sent without cash being received or that cash is sent without collateral being received. At the initiation of a trade, this is accomplished as follows. The repo buyer sends cash and the repo seller sends collateral not to each other, but to a tri-party repo agent. The agent, after verifying receipt of both cash and collateral, moves the cash into the seller's account and the collateral into the buyer's account. Similarly, at the expiration of the repo, the agent verifies that appropriate amounts of cash and collateral are available and, only then, moves the cash into the buyer's account and the collateral into the seller's account.

A second significant tri-party service ensures that repo margin is enforced along the lines described earlier in the chapter. More specifically, for every transaction, the tri-party agent verifies that the collateral received is acceptable relative to parameters set by the repo buyer, values that collateral at current market prices, and issues daily variation margin calls as required by the haircut schedule of the repo buyer.

The third significant tri-party service is to optimize the use of general collateral. As mentioned earlier, each repo buyer accepts collateral that conforms with, for example, customized constraints on asset class, credit rating, and maturity. At the same time, each repo seller has a portfolio of diverse

TABLE 10.1 Structure of the US Repo Market, as of July 2021, in $Billions.

	Tri-Party	Bilateral
Uncleared	2,964	??
Cleared	238 (GCF)	1,181 (DVP)

Sources: Federal Reserve Bank of New York; Office of Financial Research; SIFMA; and Author Calculations.

[5]Tri-party entries in Table 10.1 are outstanding collateral values; DVP entries are outstanding loan amounts.

securities it wants to post as collateral. Furthermore, not surprisingly, repo loans earn a higher rate when extended against less desirable collateral. The tri-party agent, then, accounting for the constraints across its customers, their repo trades on a particular day, and current market prices and repo rates, suggests to each repo seller an optimal allocation of its available collateral, that is, a schedule of which collateral should be posted to which repo buyer.

Table 10.1 shows that the tri-party repo market is divided into two segments. The larger segment, with nearly $3 trillion of outstanding collateral value, consists predominantly of GC trades between repo investors (e.g., money market funds) and dealers, who fund their securities positions with repo. The smaller segment, at $238 billion, is the *General Collateral Finance* or *GCF* repo service. In this "blind-brokered" market, dealers trade anonymously with each other to redistribute funds received from repo investors. GCF repo trades are *cleared* by the Fixed Income Clearing Corporation (FICC), which means that each counterparty to a GCF repo trade faces the FICC, instead of its original trade counterparty.[6] This segment of the market also includes *sponsored repo*, in which non-member repo counterparties that are backed by members can trade GCF repo on their own accounts.[7]

The bilateral repo market is summarized in the rightmost column of Table 10.1. Because tri-party is particularly focused on managing the collateral of GC trades, the bilateral market is used for the GC trades of those who, for various reasons, do not participate in tri-party, and for specials trades. Like the tri-party market, the bilateral market is divided into an uncleared and a cleared segment. Relatively little aggregated data is available about the uncleared segment precisely because its transactions are completely managed by the individual counterparties. The bilateral, cleared market comprises the *delivery-for-payment* or *DVP* service of FICC. The trades in this inter-dealer market can be arranged or blind-brokered, and they are cleared by the FICC. The size of the DVP market is about $1.2 trillion in loan value.

Figure 10.7 considers the repo market from the perspective of various sectors of market participants. Note that, in accounting terms, lending cash or buying repo is an asset, and borrowing cash or selling repo is a liability. While broker–dealers do sell repo to finance security holdings, they also run

[6]Note that clearing is much stronger than tri-party guarantees, which ensure that *if* the repo seller sends collateral and the repo buyer sends cash, *then* the collateral and cash will be exchanged.

[7]One advantage of sponsored repo is the reduction of the size of broker–dealer balance sheets. If a broker–dealer sells repo to a money market fund and then buys repo from an asset manager, the broker–dealer incurs a liability and an asset. If the broker–dealer sponsors the fund and asset manager, however, so that they can trade with each other, their trades do not appear on the broker–dealer's balance sheet. The costs and risks of sponsorship, however, may have negative accounting implications.

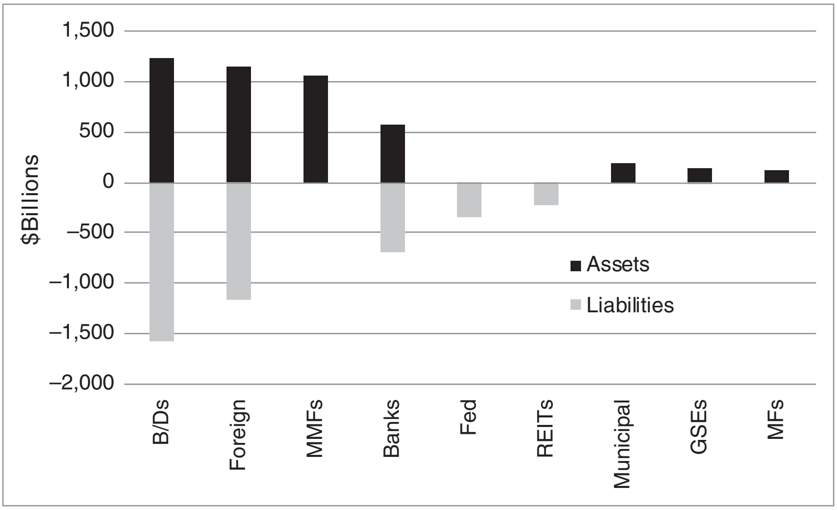

FIGURE 10.7 US Repo Market as Assets and Liabilities, by Sector, as of March 2021. *Source:* Financial Accounts of the United States, Board of Governors of the Federal Reserve System.

large *matched books* through which they make markets in repo, that is, they buy repo from some customers and sell to others in a balanced way. The combination of a matched book and long financing leads to the profile shown in the figure, namely, a large quantity of repo assets offset by repo liabilities and an additional quantity of repo liabilities. Money market funds, as discussed, are in the business of investing short-term funds and use repo to do so. The foreign sector (excluding foreign bank offices in the United States) is approximately balanced in repo assets and liabilities, while bank repo liabilities are somewhat greater than their repo assets. The Federal Reserve borrows cash in the repo market through its reverse-repo facility, as explained in Chapter O. Real Estate Investment Trusts (REITs) borrow money through repo, while the remaining sectors shown – municipals; GSEs, like FNMA and FHLMC; and mutual funds (MFs) – invest in repo as a liquid, short-term investment.

Continuing the discussion of repo market structure, Table 10.2 gives the composition of collateral in tri-party repo, ex-GCF. Underlining the safety of repo transactions, a large portion of collateral is government guaranteed, with 62% in Treasuries and 21% in government-guaranteed mortgage-backed securities (MBS). Furthermore, although not shown in the table, the $238 billion of GCF repo consists entirely of Treasury, Agency, and GSE MBS. Some of the tri-party market, however, as shown in Table 10.2, does include repo using riskier collateral.

To conclude this section, Table 10.3 gives the median haircut on collateral in tri-party, ex-GCF. Recalling that a cash lender loses money only

TABLE 10.2 Collateral Composition of Tri-Party Repo (ex-GCF), as of July 2021. Collateral Value Is in $Billions.

Asset Class	Collateral Value	Percent
Treasuries	1,847	62.3
Agency and GSE MBS	627	21.1
Equities	209	7.1
Corporates, Investment Grade	74	2.5
Agency and GSE CMOs	52	1.8
Corporates, High Yield	42	1.4
Agency and GSE Debt	33	1.1
Other	80	2.7
Total	2,964	100

Sources: Federal Reserve Bank of New York; SIFMA; and Author Calculations.

TABLE 10.3 Median Tri-Party Repo Haircuts (ex-GCF), as of July 2021. Haircuts Are in Percent.

Asset Class	Haircut
Treasuries; Agency Debt and MBS	2
Agency and GSE CMOs	4
Corporates (IG); Money Mkt; Private CMO (IG); Int'l	5
ABS (IG)	7
Equities; Corporates (HY); ABS(HY); Private CMO (HY); Munis	8

Sources: Federal Reserve Bank of New York; and SIFMA.

if its counterparty defaults and the collateral falls in value by more than the haircut, cash lenders demand higher haircuts for riskier collateral. Risk in this context is measured not only by price volatility and the possibilities of large price declines but also by illiquidity: in the event of a default, lenders recover only what they can realize through the sale of the collateral they hold. Consistent with these considerations, the median haircuts in Table 10.3 do increase with risk. The most liquid, government-guaranteed collateral requires a haircut of only 2%, while the riskiest and least liquid securities require a haircut of 8%.

10.4 SOFR

In the transition away from LIBOR in the United States, discussed in detail in Chapter 12, regulators and others have advocated for a transition to the *Secured Overnight Financing Rate*, or *SOFR*. Intended to represent the rate

TABLE 10.4 SOFR and Treasury Repo Rates, as of May 14, 2021. Rates Are in Basis Points.

1st %ile	25th %ile	Median/ SOFR	75th %ile	99th %ile	Volume ($billions)
–4	–1	1	1	15	865

Source: Federal Reserve Bank of New York.

of secured, overnight borrowing, the Federal Reserve Bank of New York calculates SOFR using daily repo transactions in the tri-party and DVP markets. However, DVP repo transactions are trimmed so as to exclude the lower-rate, presumably specials trades, which reflect the idiosyncrasies of lending on individual bond issues. In particular, the 25% of DVP repo trade volume with the lowest rates are excluded from the data used to calculate SOFR.

Given all included transactions on a given day, SOFR that day is calculated as the "volume-weighted median" rate. This means that SOFR is determined such that 50% of the volume of loan amounts are at a lower rate and 50% at a higher rate. Table 10.4 provides information about the calculation of SOFR on May 14, 2021. From $865 billion of trades on that day, SOFR is set to the volume-weighted median rate, which is one basis point. Trades representing 50% of the volume – from the 25th to the 75th percentiles – include rates between minus one and plus one basis point.

10.5 GC AND SPECIAL REPO RATES

This section discusses the interest rates on repo loans. Conceptually, there is a distinct rate for every repo term and for every type of collateral: one rate for overnight repo on investment-grade corporate collateral; another rate for one-month repo on government-guaranteed mortgage-backed securities; etc. There is, however, a benchmark repo rate, called "the GC rate," which is the overnight rate earned by lenders willing to accept any Treasury collateral.

Some insight into the determination of the GC rate can be gained by comparing the effective federal funds rate (EFFR) with the GC rate. The interbank EFFR is a highly creditworthy, overnight, unsecured rate, which can be compared with the GC rate, a highly creditworthy, overnight, secured rate. To this end, Table 10.5 shows percentiles of the differences between effective fed funds and the GC rate, from January 2001 to July 2021.

The median, or 50th percentile, of the difference is one basis point. It is not surprising that an overnight unsecured rate, EFFR, tends to be above an overnight secured rate, GC. It is also not surprising that the spread is relatively small, because, as an overnight rate between banks, EFFR is a rate on a very high-quality loan. What is surprising and instructive, however,

TABLE 10.5 Percentiles of Daily Spreads of the Federal Funds Effective Rate over the General Collateral Treasury Repo Rate, from January 2, 2001, to July 12, 2021. Spreads Are in Basis Points.

Percentile	Spread
0.5	−37.4
5.0	−13.0
25.0	−4.0
50.0	1.0
75.0	5.0
95.0	15.0
99.5	112.6

Sources: Federal Reserve Bank of New York; Barclays Capital; and Author Calculations.

is by how much these two rates can diverge, in either direction. The 95th percentile of the spread, for example, has EFFR 15 basis points above GC, while the 5th percentile has EFFR 13 basis points below GC. And at greater extremes, for the 0.5 and 99.5 percentiles, EFFR is 113 basis points above and 37 basis points below GC, respectively.

To explore these large deviations, Figure 10.8 graphs the EFFR–GC spread over time. There are five particularly large and sudden increases in the GC rate, which result in large, negative spreads. In general, these episodes are all manifestations in the Treasury repo market of a relatively desperate need to raise cash and a willingness to pay high rates to do so. The first was in September 2001, shortly after the terrorist attack on the World Trade Center. Damage to infrastructure in downtown New York City temporarily disrupted the normal flow of funds and sent market participants, including banks, scrambling for cash. Liquidity provision by the Federal Reserve and infrastructure repair restored market order in a couple of weeks.

The next events, September 30, 2016, and December 31, 2018, are at quarter end and year end, respectively. Many market participants are believed to engage in "window dressing" at the end of reporting periods, possibly to impress investors and possibly to meet various regulatory requirements. This entails temporary measures such as increasing liquidity and cutting back on loans, which, if prevalent enough, push up repo rates as borrowers hunt for cash among willing lenders. Quarter- and year-end market disruptions have become less frequent over the decades, as the Federal Reserve has successfully anticipated and accommodated the resulting scarcities of cash. As the two highlighted events suggest, however, quarter- and year-end disruptions are not extinct.

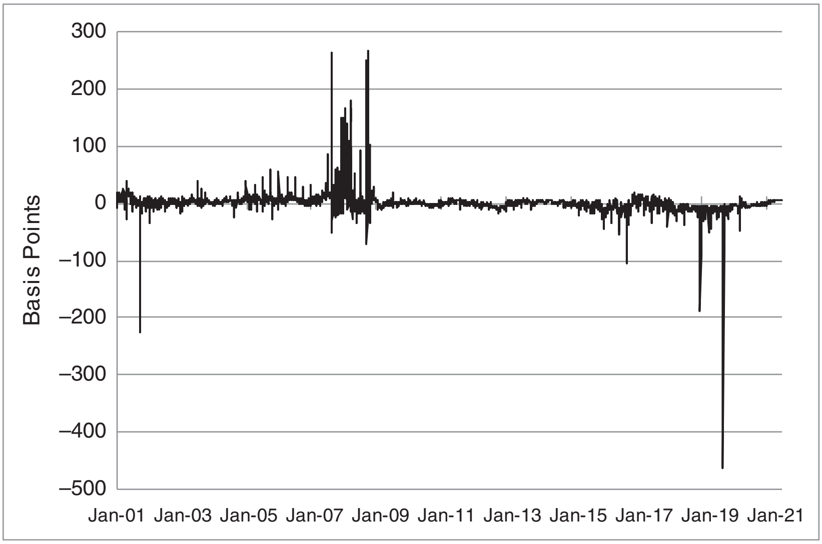

FIGURE 10.8 Daily Spreads of the Federal Funds Effective Rate over the General Collateral Treasury Repo Rate, from January 2, 2001, to July 12, 2021. *Sources:* Federal Reserve Bank of New York; Barclays Capital; and Author Calculations.

Continuing with the next negative spread events in Figure 10.8, in September 2019, a number of cash-draining incidents happened at about the same time: quarterly corporate tax payments; settlement of newly issued Treasuries; and a holiday in Japan, which temporarily sidelined some Japanese investors in US repo. But the magnitudes of these incidents seem small relative to the realized cash scarcity and to the Federal Reserve's subsequent large and long-lasting injections of liquidity. Furthermore, banks did not lend to take advantage of the resulting abnormally high repo rates despite having significant reserves. Another, though smaller, negative spread event occurred in March 2020, when banks also refrained from lending at seemingly advantageous rates and from purchasing seemingly mispriced Treasuries, which was the larger dislocation at the time. The reluctance of banks to commit funds during these market dislocations is discussed later in the chapter.

In addition to large negative spreads, Figure 10.8 shows several instances of very large positive spreads, particularly through the financial crisis of 2007–2009. The two largest such events were in the second half of August 2007, which marked the beginning of the subprime mortgage crisis, and mid-September 2008, when Lehman Brothers filed its bankruptcy. During credit events, when market participants worry about each other's ability to satisfy their obligations, there is a rush out of risky assets and into safe assets, including GC repo. As a result, GC repo rates decline and the EFFR-GC spreads become large.

TABLE 10.6 Treasury Special Spreads, as of May 27, 2021. Entries Are in Basis Points.

	2yr	3yr	5yr	7yr	10yr	20yr	30yr
On-the-Run	5	1	13	9	3	2	4
Old	2	2	2	1	11	2	12
Double-Old	2	10	6	8	10	2	3
Triple-Old	2	12	2	3	3	2	2

Source: Citi.

Having completed an overview of the GC rate, the discussion turns to special rates, the rates earned on loans when particular bonds are required as collateral, usually to facilitate short sales. While any bond can "trade special," that is, be required as collateral in a special repo trade, the most common specials in the US repo market are recently issued Treasuries bonds. As mentioned in earlier chapters, the most recently issued Treasuries tend to be the most liquid, and, for this reason, often trade at a premium to otherwise similar bonds. Similarly, market participants who want to short bonds prefer to trade in these bonds and are willing to pay to do so. And the way they pay more to short these bonds is to lend at a relatively low repo rate when taking these bonds as collateral.

The special spread of a bond to a given term is defined as the difference between the GC rate and the special rate on that bond to that term. Table 10.6 shows the overnight special spreads of various recently issued bonds. The *on-the-run (OTR)* bond of each maturity is the most recently issued bond of that maturity, the *old* bond is the second most recently issued, and so forth. The overnight special rate for the on-the-run five-year bond, for example, is 13 basis points below the GC rate, while the old 10-year special rate is 11 basis points below GC. The special spread for each bond is idiosyncratic and time varying, depending on supply and demand for that bond as collateral, which, by the way, often involves a different set of market participants than the supply and demand for sale or purchase of the bond outright. Despite the idiosyncratic nature of special spreads, more can be said about them in terms of the Treasury auction cycle.

Figure 10.9 graphs the overnight special spread of the OTR 10-year Treasury note over time. Recall from Chapter O that a new 10-year note is issued every February, May, August, and November and is reopened after one month and again after two months. The special spreads in the figure, therefore, correspond to different bonds as new 10-year notes are issued. The dotted, vertical, gray lines mark the dates of these new issues; the solid, gray lines mark their first reopenings; and the dashed, gray lines mark their second reopenings. The first lesson from the figure is that, as mentioned in the previous paragraph, the behavior of the special spread is very

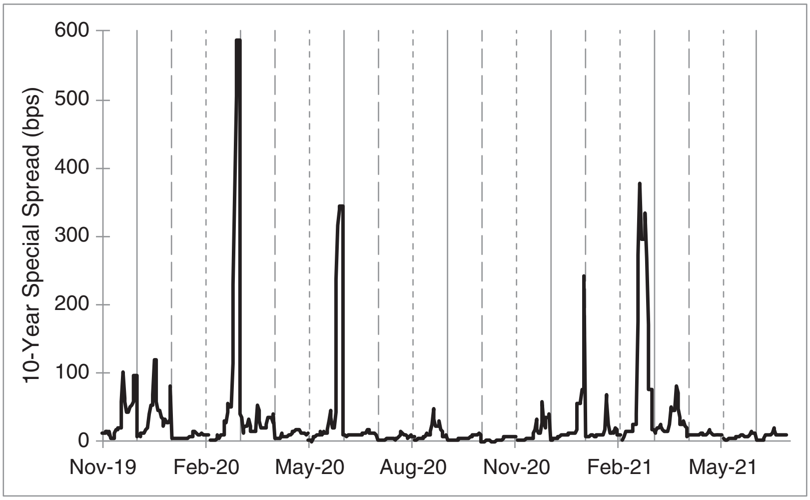

FIGURE 10.9 On-the-Run 10-Year Treasury Special Spread. *Sources:* Barclays Capital; and Author Calculations.

idiosyncratic: some OTR 10-year notes become very special, and some do not.[8] The second lesson is that the OTR notes tend to become particularly special into their first reopening. Of the bonds shown in the graph, only one clearly violates this rule: its specialness peaks in January 2021, which corresponds to its second reopening.

Figure 10.10 describes the average behavior of special spreads over the auction cycle from November 2008, when the Treasury auction cycle first settled into quarterly issuance and monthly reopenings, through July 2021. This period of time spans the issuance of 51 10-year notes. The thick, black line in the figure gives the average special spread across these 51 notes as a function of the number of days to the first reopening. For example, with 13 days to the first reopening, the average special spread across the 10-year notes in the sample is 85 basis points. The thick, dark gray line in the figure gives the average special spread as a function of the days to the second reopening, and the thick, light gray line gives the average special spread as a function of the days to the next new issue.[9]

---

[8]The extremely large special spreads in March 2020 actually entail negative special rates. Since May 2009, there has been a penalty for failing to deliver bonds that had been sold at a rate equal to the maximum of zero and the difference between 3% and the fed funds rate. Therefore, traders can be willing to lend cash at negative rates to borrow bonds so as to avoid failing to deliver and so as to avoid the penalty.

[9]More precisely, the number of days is to the settlement date of the bonds sold at the first reopening, second reopening, or new issue.

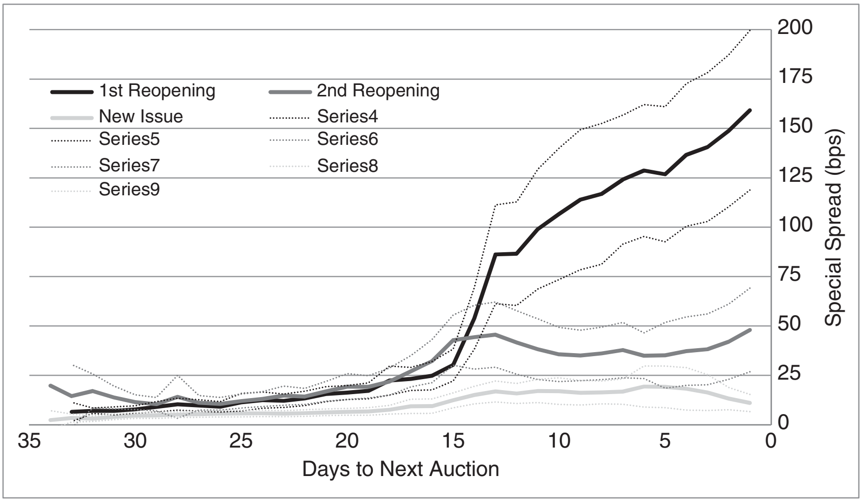

FIGURE 10.10 Average 10-Year Treasury Special Spread over the Auction Cycle, with Two Standard Deviation Confidence Intervals, from November 2008 to July 2021. *Sources:* Barclays Capital; and Author Calculations.

The figure shows that, on average, the special spread is relatively low right after an auction, but then increases into the next auction. Furthermore, the increase into the first reopening is highest, reaching 156 basis points the day before the first reopening. The increase into the second reopening is smaller, reaching 48 basis points the day before the second reopening. And the increase into the new issue is smallest, reaching less than 20 basis points in the days before the new issue.

The market behavior behind these results is as follows. When a note is first issued, the previously issued note is still the most liquid in the maturity range and the most popular choice as a short. As the first reopening approaches, two things happen. First, liquidity shifts from the previously issued note to the new OTR note. Second, dealers increase shorts in the OTR note to hedge their market-making purchases of notes issued in the next auction. This segment of short demand disappears just after the first reopening but builds up again into the second reopening. The special spread does not increase as much as into the first reopening, however, because the supply of the OTR notes is now significantly larger. The cycle then repeats once more, into the issue of the next OTR note, but the supply of the current OTR notes is now too large for the special spread to increase by very much.

While the story in the previous paragraph describes average behavior, Figure 10.9 illustrates how the behavior of special spreads varies across bond issues. Figure 10.10 emphasizes this point by showing, with dotted lines around the averages, two-standard-deviation confidence intervals. For example, while the average special spread of an OTR 10-year one day before its first reopening is 156 basis points, the two-standard-deviation range around that average is from 116 to 196 basis points.

To conclude the discussion of special spreads, recently issued bonds can command a price premium not only because they are relatively liquid but also because they have value as collateral in the repo market. Put another way, not only can holders of recent issues trade out of their holdings easily, but they can borrow money cheaply by posting these issues as collateral. In fact, along the lines of Figure 10.10, traders can model the special spread of an individual issue over time and thereby estimate the value, over its life, which arises from its expected repo specialness.

10.6 LIQUIDITY MANAGEMENT AND CURRENT REGULATORY ISSUES

A financial institution can borrow funds in many ways, some of which are more stable than others, that is, some of which can be easily maintained under financial stress and some of which cannot. The most stable source of funds is equity capital, because equity holders do not have to be paid on any schedule and because they cannot compel a redemption of their shares. Slightly less stable is long-term debt, because bondholders are paid interest and principal as set out in bond indentures. At the other extreme of funding stability is short-term unsecured funding, like commercial paper: these borrowings have to be repaid in a matter of weeks or months, as they mature, when the institutions, under adverse conditions, might not be able to borrow money elsewhere. Not surprisingly, the more stable sources of funds are usually more expensive in terms of the expected return required by the providers of funds. Through *liquidity management*, firms balance the costs of funding against the risks of finding themselves without enough funding to survive.

In the spectrum of financing choices, repo borrowing is cheap, but relatively unstable. This instability is evident with respect to short-term repo on credit-sensitive and relatively illiquid assets. In times of financial stress, many market participants sell these low-quality assets, both to reduce risk and to raise liquidity. Meanwhile, in the repo market, lenders both raise haircuts and reduce lending on low-quality collateral. This makes it difficult for repo sellers of that collateral – like trading desks funding inventory or hedge funds leveraging trades – to raise cash. And if it turns out that they cannot find sufficient funding, they might be forced to unload their holdings in fire sales that result in significant losses.

The instability of repo borrowing on high-quality collateral, like Treasuries, is less evident. Before the financial crisis of 2007–2009, it was widely believed that a firm could always fund its Treasuries. First, in times of stress, lenders are looking for liquid and safe investments. Second, even if the creditworthiness of the cash borrower is called into question, the cash lender has Treasury collateral, which, if necessary, can be liquidated relatively easily. The financial crisis dispelled this complacency. Lenders sought to minimize their connections to troubled financial institutions, including their buying of Treasury repo. While they could, indeed, legally liquidate those Treasuries in the event of a default, the timing and manner of their doing so could be questioned in bankruptcy proceedings and could drag them into years of litigation.

The financial crisis of 2007–2009 found many financial institutions with highly leveraged positions in mortgage-related assets. As these assets lost value, an overreliance on wholesale funding,[10] from various forms of commercial paper to repo, accelerated the onset of distress. In response, the bank regulatory regime has changed in ways that significantly impact repo participants and markets.

First, banks have become subject to liquidity ratios, namely, the *liquidity coverage ratio* (LCR) and the *net stable funding ratio* (NSFR). The LCR requires that banks hold enough high-quality, liquid assets (HQLA) to meet net cash outflows over a 30-day stress scenario. Reserves held in the Federal Reserve System and Treasuries receive full credit as HQLA, while lower-quality assets may fulfill part of the requirement at a haircut. The NSFR requires that bank funding of relatively illiquid assets be stable over a one-year horizon. The direct and intended effect of these ratios is to limit the amount of short-term, wholesale funding, including repo. Less directly and perhaps less intended, regulators' implementation of the LCR, combined with their additional stress tests of bank liquidity, have pushed banks to rely heavily on reserves to satisfy HQLA requirements.[11] This implies that regulation, as currently implemented, discourages using reserves to lend money and take Treasury collateral.

Second, banks have become subject to a leverage ratio, called the *supplementary leverage ratio* (SLR) in the United States. Unlike risk-based capital requirements, which require banks to hold capital in proportion to the perceived risks of their assets, the leverage ratio requires that bank capital exceed a certain percentage of the value of all assets, regardless of risk. But penalizing large positions regardless of risk discourages large repo positions, which have very little risk per dollar and commensurately low return per dollar. Consider, for example, a matched book business in Treasury repo,

[10]The term *wholesale funding* refers to funding provided by institutional, professional, and particularly sophisticated investors, in contrast with retail funding, like deposits.

[11]See, for example, Nelson, B., and Covas, F., (2019), "Bank Regulations and Turmoil in Repo Markets," Bank Policy Institute, September 26.

mentioned previously, which lends and borrows money on Treasury collateral. This market-making business in extremely safe transactions can make money and cover fixed costs only in large size, but the leverage ratio discourages size.

These two sets of rules may have been partly responsible for the aforementioned September 2019 and March 2020 episodes in which banks did not use their abundant reserves to lend in the Treasury repo market at abnormally high rates. Or, put another way, regulators are still calibrating the necessary quantity of reserves in the context of a relatively new regulatory regime. In fact, one of the ways the Federal Reserve and other regulators addressed the March 2020 disruption was to exclude reserves and Treasuries from the SLR temporarily, from May 15, 2020, to March 31, 2021.

10.7 CASE STUDY: MF GLOBAL'S REPO-TO-MATURITY TRADES

In March 2010, MF Global was a broker–dealer and futures commission merchant, in the business of making markets in exchange-traded and over-the-counter derivatives.[12] Experiencing poor performance, in particular, a fiscal-year loss of $195 million in its core businesses, the company hired Jon Corzine as CEO. Corzine worked for Goldman Sachs from 1976 to 1999 and was chair and CEO from 1994 to 1999. He was a US senator from New Jersey from 2001 to 2006 and the governor of New Jersey from 2006–2010.

Corzine was told by the ratings agencies that MF Global would need to increase earnings to avoid a ratings downgrade, with Moody's specifying a target of between $200 million and $300 million pretax annual profits by September 30, 2011. Faced with this timeline, and pursuing part of a larger vision for the company, Corzine focused on proprietary trading and discovered a particularly attractive opportunity in European government bonds.

Europe was in the throes of its sovereign debt crisis, with rising concerns about the solvency of countries in weaker fiscal positions. As one way of dealing with the situation, the European Union created the European Financial Stability Facility (EFSF), in which all of the countries in Europe, as of May 2010, essentially guaranteed each other's debt until the end of June 2013.

[12]This account draws heavily on Skyrm, S. (2013), *The Money Noose*, Brick Tower Press; and Freeh, L. (2013), "Report and Investigation of Louis J. Freeh," United States Bankruptcy Court for the Southern District of New York. The case also appears in the broader context of liquidity risk management in Tuckman, B. (2017), "Survive the Droughts, I Wish You Well: Principles and Cases of Liquidity Risk Management," *Financial Markets, Institutions & Instruments*.

TABLE 10.7 European Short-Term Government Bond Rates, as of December 2010. Rates Are in Percent.

Country	Short-Term Bond Rate
Germany	0.70
Italy	2.25
Spain	3.32
Portugal	4.50
Ireland	6.00

In this setting, Corzine proposed repo-to-maturity (RTM) trades in the short-term bonds of the weaker European sovereigns. More specifically, MF Global would buy the short-term debt of the weaker sovereigns and sell repo to the maturity of the bonds. Table 10.7 gives the prevailing rates on short-term government bonds at that time. Term repo rates on these bonds were essentially risk-free rates, and approximately equal to Germany's short-term bond rate. Furthermore, haircuts on European sovereign repo were about 3%. Hence, in these RTM trades, MF Global could earn the rates in Table 10.7 on its bond purchases, funded at a repo interest rate cost of about 0.70%.

These RTM trades were designed to minimize risk. First, no investments were to be made in the government bonds of Greece, the weakest sovereign at the time. Second, by investing in short-term bonds only, specifically those that mature before the earliest possible expiration of the EFSF in June 2013, default risk was minimal: bondholders would lose principal only if Europe as a whole reneged on its EFSF commitments, which was considered extremely unlikely. Third, by selling the repo to the maturity of the bonds, there was no risk of losing funding in the middle of the trade.[13] Funding the trades with overnight repo, by contrast, runs the risk that funders refuse to roll as sovereign creditworthiness deteriorates and prices fall. In that scenario, instead of realizing the full principal amount at maturity, MF Global could be forced into selling the bonds early at a loss in order to pay off the repo loans.

The RTM trades were also attractive from an accounting perspective. Assume that MF Global owns a bond today. Then, by selling the repo today, the firm receives money today and extinguishes future cash flows: the bond's principal payment at maturity is used to pay off the repo loan. By this reasoning, accounting rules at the time regarded selling the repo as equivalent to selling the bond today. Therefore, when MF Global bought a bond and

[13]The repo term was actually two days before maturity, a detail not discussed further in the text.

sold the repo, it could immediately book the profits from the trade (i.e., the difference between the bond coupon and repo interest) and did not have to disclose the trades to investors. The immediate booking of profit was particularly appealing to Corzine, who was under pressure to show earnings to the rating agencies. But the failure to disclose the trades would contribute to the firm's problems later on. In any case, MF Global quickly grew its position in RTM trades, from which it booked $15 million in profits in the third quarter of 2010; $47 million in the fourth quarter; $25 million in the first quarter of 2011; and more than $38 million in the second quarter.

MF Global's risk manager, Michael Roseman, understood one significant risk of the RTM trades. Had the repo counterparty been a dealer, the haircut would likely have been fixed over the term of the repo. There would be variation margin calls, of course, but those are limited to restoring the initial haircut. MF Global's repo, however, was with the London Clearing House (LCH), which has the right to change haircuts at any time. Roseman recognized that, if the sovereign credit situation worsened and LCH increased haircuts, MF Global might not have enough cash on hand to stay in the trades. That is, bonds might have to be sold at a loss to raise cash to post as additional margin. He therefore supported a position limit of $4.75 billion, because, at that size and at an assumed worst-case scenario of average haircuts at 15%, MF Global would be able to meet future margin requirements with readily available cash.

MF Global grew its RTM positions rapidly, reaching the $4.75 billion limit by year-end 2010. MF Global wanted to increase its position further, however, and fired Roseman at the end of January 2011. The trajectory of MF Global's position is shown in Figure 10.11. Positions increased dramatically in 2011, reaching a peak of $11.5 billion in June. This peak position included a significant volume of reverse RTM trades, in which MF Global shorted French government bonds and loaned the proceeds in repo. These reverse RTM trades were expected to reduce margin requirements and the volatility of variation margin calls, because, to the extent that the prices of weaker sovereign bonds and French bonds move together, changes in the collateral value of the weaker sovereigns posted by MF Global are offset by changes in the collateral value of French bonds held by MF Global. To the extent that weaker sovereign bond prices fall more than French bonds, of course, the reverse RTM trades do not hedge the RTM trades. And, in any case, the reverse RTM trades do not reduce the risk of repo haircuts increasing for the weaker sovereigns.

Along with MF Global's position size, Figure 10.11 shows haircuts over time for repo on the bonds of two of the more challenged European sovereigns, namely, Ireland (IE) and Portugal (PT). Haircuts were increased several times, from 3%, when MF Global first put on its trades, to 80% in July 2011. In other words, for these sovereigns, required margin on a repo of 100 of collateral value rose from 3 to 80. These dramatic increases

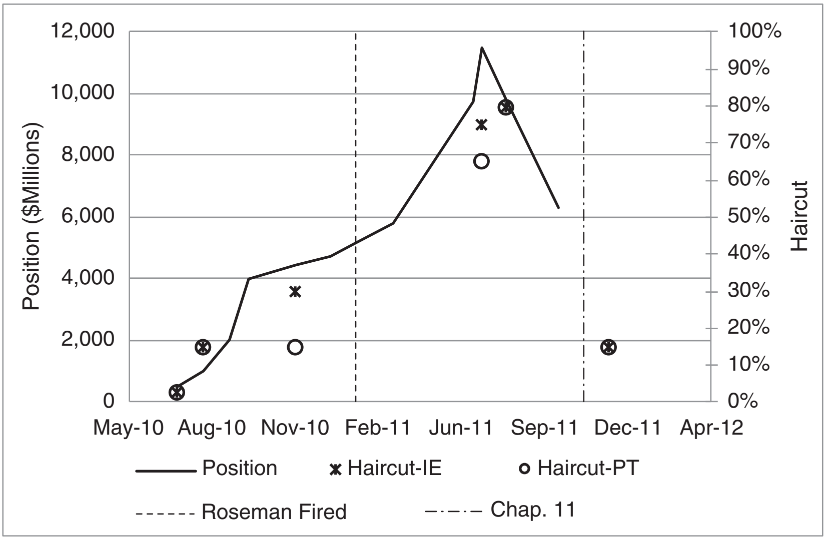

FIGURE 10.11 MF Global Repo-to-Maturity Positions and Margin.

TABLE 10.8 MF Global Repo-to-Maturity Positions and Margin.

Date	Position ($billions)	Margin ($millions)	Margin (%)
06/06/11	>=5.8	170	2.9
06/20/11	11.5	550	4.8
09/30/11	6.3	417	6.6
10/27/11	6.3	665	10.6

Source: Skyrm, S. (2013), *The Money Noose*, Brick Tower Press.

in margin were likely due partially to deteriorating sovereign credits and partially to the size and perceived precariousness of MF Global's position.

The cash implications for MF Global are summarized in Table 10.8, which shows – across all sovereigns – RTM positions, dollar margin, and percent margin over time. As position size increased to its peak in June 2011, the percent margin rose somewhat from its initial value. From then on, however, as LCH raised haircuts on the weaker sovereigns, the percent margin kept increasing despite MF Global's having reduced positions dramatically. Correspondingly, the dollar margin required for a position of about $6 billion increased from $170 million in early June to $665 million in late October. The table also shows that MF Global would have done well to stay within Roseman's position limit, which was sustainable even if average haircuts increased to 15%.

In the second and third quarters of 2011, problems in addition to the cash drain from rising haircuts caused counterparties and creditors to withdraw support from MF Global. First, regulators and accountants decided that selling repo did not constitute the sale of a bond.[14] This change led regulators to increase MF Global's capital requirements significantly and forced MF Global to disclose its RTM positions, which, in the midst of the European sovereign debt crisis, surprised the market unfavorably. Second, MF Global lost $192 million in its core businesses, that is, unrelated to the RTM trades. MF Global was downgraded to below investment grade, and it filed for Chapter 11 at the end of October 2011.

MF Global lost over $400 million in the liquidation of its RTM positions, a huge number relative to the profits given earlier. Particularly provoking is the fact that 50% of the RTM portfolio, and, for example, over 90% of its position in Italian sovereign bonds, matured within two months of the bankruptcy. And, as shown in Figure 10.11, haircuts on Ireland and Portugal sovereign bond repo fell back to 15% by the end of November. In short, had MF Global sized its position to account for the financing risk of increased haircuts, it might have profited rather than have been destroyed by its RTM trades.

[14]This change was partly because the repo expired two days before the expiration of the bond, as mentioned in footnote 13.

CHAPTER 11

Note and Bond Futures

Futures on government bonds are among the most liquid fixed income products around the world, used to hedge interest rate risk and to take positions on changes in bond prices. Figure 11.1 shows how the total DV01 traded across all US Treasury bonds and futures is divided across instruments.[1] By this metric, the 10-year note futures contract is the single instrument with the greatest volume, and volumes traded in other futures contracts rival those of on-the-run Treasury securities. Futures contracts are appealing because of their liquidity, and also because relatively little cash is required to establish sizable positions. The US Treasury futures discussed in this chapter trade on the CBOT (Chicago Board of Trade), which is part of the CME Group.

A *forward* contract on a bond is an agreement that fixes the price at which a bond is to be bought and sold on some future date. Forward contracts on bonds are rarely traded in the United States, because, as is shown next, the same economics can be achieved by trading bonds and repo. Understanding forward contracts is important, however, because futures are essentially forwards with the complexities of *daily settlement* and various *delivery options*. This chapter, therefore, first describes forward contracts and then adds these features of futures contracts. The exposition includes a discussion of *basis trades*, in which futures are traded against synthetic forward bond positions, and the chapter concludes with a case study about basis trades through the volatility of March 2020, at the start of the COVID pandemic and economic shutdowns.

For concreteness in explaining concepts, this chapter focuses on the 10-year US Treasury note contract. Other US bond futures are similar, as are many government bond futures around the world. Contracts on UK and Chinese government bonds, for example, are particularly similar to those

[1]This graph is Figure 1 of Baker, L., McPhail, L., and Tuckman, B. (2018), "The Liquidity Hierarchy in the US Treasury Market," Commodity Futures Trading Commission, December 3.

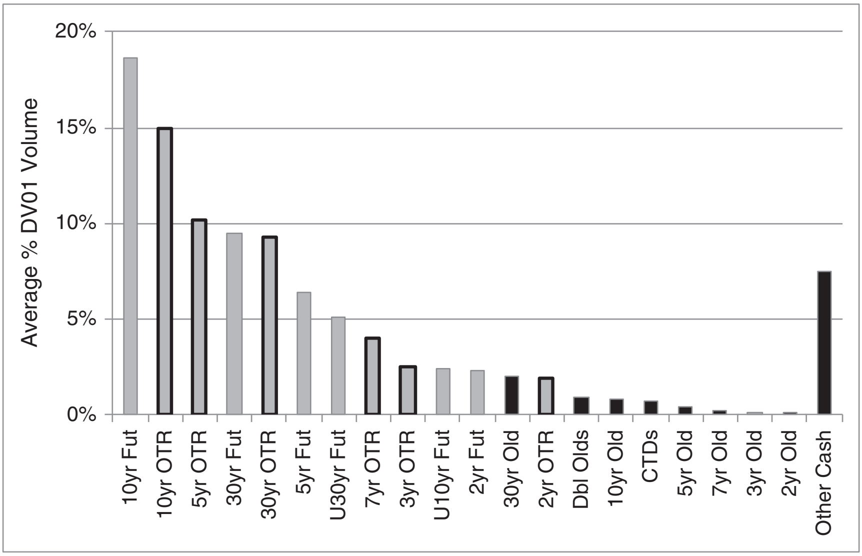

FIGURE 11.1 Percent of Total DV01 Traded in US Treasury Bonds and in US Treasury Note and Bond Futures, by Instrument, from July 10, 2017, to June 1, 2018. "Fut" Denotes Futures Contracts on Bonds of Various Maturities; and "U30yr Fut" and "U10yr Fut" Are the "Ultra" 30- and 10-year Futures Contracts. "OTR" Denotes the On-the-Run or Most Recently Issued Treasury Bond at a Given Maturity; "Old" Denotes the Next Most Recent; and "Dbl Old" Denotes the Most Recent After the Old. "CTDs" Denote the Cheapest-to-Deliver Bonds into Futures Contracts. "Other Cash" Denotes All Other Treasury Bonds.

in the United States in that they embed both *quality* and *timing* options. Contracts on government bonds in Europe and Japan are also similar, but simpler, in that they do not have a timing option.

Appendix A11.2 outlines the pricing of forward and futures contracts on bonds in a term structure model.

11.1 FORWARD CONTRACTS AND FORWARD PRICES

In a *forward* bond contract, the counterparties agree on a price at which to trade a bond at some time in the future. Consider a forward contract, traded on May 14, 2021, to purchase $100,000 face amount of the US Treasury 2.875s of 05/15/2028 for 109.72 on September 30, 2021. In this example, the initiation or trade date is May 14, 2021; the *underlying security* is the

TABLE 11.1 A Forward Contract on the US Treasury 2.875s of 05/15/2028.

Underlying Security	2.875s of 05/15/2028
Principal Amount	$100,000
Trade Date	05/14/2021
Spot Settlement Date	05/17/2021
Spot Price	110.77344
Accrued Interest, Spot Settlement (2/184 days)	0.01562
Forward Settlement Date	09/30/2021
Accrued Interest, Forward Settlement (138/184 days)	1.07813
Repo Rate to Forward Settlement Date	0.015%
Days from Spot to Forward Settlement	136
Forward Price	109.71721

2.875s of 05/15/2028; the *forward date*, *expiration date*, *delivery date*, or *maturity date* is September 30, 2021; and the *forward price* is 109.72. The counterparty committing to purchase the bond on the forward date is the *buyer* of the contract and is *long* the forward, while the counterparty on the other side is the *seller* of the contract and is *short* the forward. Table 11.1 summarizes the terms of this forward trade and gives other data that are used presently.

By definition, the forward price is such that the buyer and seller are willing to enter into the forward agreement without any exchange of cash. This implies that the initial value of the forward contract is zero. Over time, however, the value of the forward position may rise or fall. Continuing with the example, say that, just after the agreement was struck, the market forward price increases to 110. The contract to buy the bond at 109.72 now has positive value to the buyer and negative value to the seller. The buyer could immediately sell the bond through a new forward contract at 110, which makes the contract worth 110 – 109.72, or 0.28, as of 09/30/2021, or the present value of 0.28 as of today. The seller of the original contract, however, having locked in a forward price of 109.72, would have to pay the buyer to exit the contract. In any case, note that the "value of a forward contract" denotes the value of an existing contract under current market conditions, while the "forward price" denotes the price at which the underlying bond is to be traded on the forward date.[2]

[2]This terminology can be confusing because it differs from that of spot settlement. Most notably, the "value of a bond" is typically used interchangeably with the "bond price."

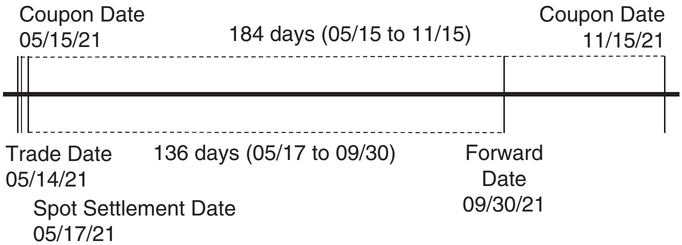

FIGURE 11.2 A Forward Contract on the US Treasury 2.875s of 05/15/2028.

Forward bond prices can be determined by arbitrage arguments. To demonstrate, consider the following strategy in the context of the example, using the data in Table 11.1 and referring to the timeline in Figure 11.2:

On May 14, 2021:

- Buy \$100,000 face amount of the 2.875s of 05/15/2028 for 110.77344 plus accrued interest of 0.01562, for settlement on May 17, 2021, for an invoice price of 110.78906 and a dollar invoice of \$110,789.06.
- Sell the repo to the forward date, that is, borrow \$110,789.06 from May 17, 2021, to September 30, 2021, at the repo rate of 0.015%, and post the 2.875s of 05/15/2028 as collateral.
- Note: no cash is generated or required on this date.

On September 30, 2021:

- Repay the repo loan, now grown to $110,789.06 \times (1 + 0.015\% \times 136/360) = \$110,795.34$.
- Take back the \$100,000 face amount of the 2.875s of 05/15/2028.
- Note: the bonds have effectively been purchased at a full price of 110.79534, which, after subtracting the accrued interest of 1.07813 as of this date, equates to a flat price of 109.71721.

This strategy buys the bond and sells the repo, which establishes a position equivalent to a long forward and, therefore, is called a *synthetic* forward position: there is no cash flow on the trade date and the bond is effectively purchased on the forward date. By arbitrage, therefore, the forward price, that is, the price at which the bond is bought through the forward contract, must equal the price at which the bond is bought through this synthetic forward, that is, 109.71721. More formally, if the forward price, F, were greater than 109.71721, an arbitrageur could sell the forward at F and buy the bond forward through the strategy at 109.71721, thus locking in a riskless profit of $F - 109.7172$ as the forward date. Similarly, if F were less than 109.71721, the arbitrageur could buy the forward and sell the bond forward through the strategy to lock in a riskless profit of 109.71721 – *F*. Hence, the only forward price consistent with no arbitrage opportunities is = 109.71721.

Algebraically, the full forward price can be written as,

$$109.71721 + 1.07813 = (110.77344 + 0.01562)\left(1 + \frac{0.015\% \times 136}{360}\right) \tag{11.1}$$

In words, the full forward price equals the future value of the full spot price to the forward delivery date at the bond's repo rate.

The derivation of the arbitrage-free forward price requires an extra step when there is an intermediate coupon payment, that is, a coupon payment between the spot and forward settlement dates. To illustrate, consider a forward contract on the 1.125s of 02/15/2031 over the same dates as the previous example, except, of course, the coupon payment dates differ. Table 11.2 and Figure 11.3 give the data for this example.

To replicate this long forward contract, implement the following strategy:

On May 14, 2021:

- Buy $100,000 face amount of the 1.125s of 02/15/2031 for 95.50781 plus accrued interest of 0.28280, for settlement on May 17, 2021, for an invoice price of 95.79061 and a dollar invoice of $95,790.61.
- Sell the repo to the forward date, that is, borrow $95,790.61 from May 17, 2021, to September 30, 2021, at the repo rate of 0.015%, and post the 1.125s of 02/15/2031 as collateral.

TABLE 11.2 A Forward Contract on the US Treasury 1.125s of 02/15/2031.

Underlying Security	1.125s of 02/15/2031
Principal Amount	$100,000
Trade Date	05/14/2021
Spot Settlement Date	05/17/2021
Spot Price	95.50781
Accrued Interest, Spot Settlement (91/181 days)	0.28280
Forward Settlement Date	09/30/2021
Accrued Interest, Forward Settlement (46/184 days)	0.14063
Repo Rate to Forward Settlement Date	0.015%
Days from Spot to Forward Settlement	136
Forward Price	95.09290

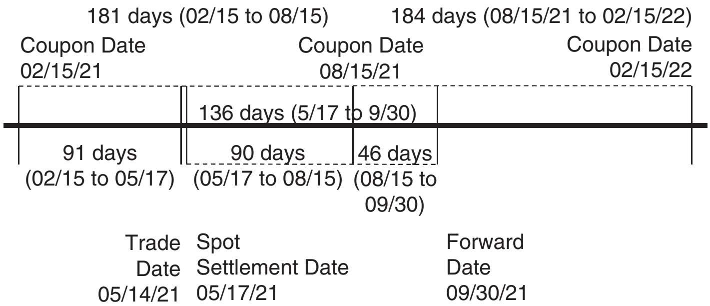

FIGURE 11.3 A Forward Contract on the US Treasury 1.125s of 02/15/2031.

- Note: no cash is generated or required on this date.

On August 15, 2021:

- Use the bond's coupon payment of $\$100{,}000 \times 1.125\%/2 = \$562.50$ to reduce the repo borrowing. More specifically, over the 90 days from May 17, 2021, to August 15, 2021, the repo loan balance grows to $\$95,790.61(1 + 0.015\% \times 90/360) = \$95,794.20$. Reducing this balance by the coupon payment leaves a balance of \$95,231.70.
- Note: no cash is generated or required on this date.

On September 30, 2021:

- Repay the repo loan, which, over the period August 15, 2021, to September 30, 2021, grows to $95,231.70 \times (1 + 0.015\% \times 46/360) = \$95,233.53$.
- Take back the \$100,000 face amount of the 1.125s of 02/15/2031.
- Note: the bonds have effectively been purchased at a full price of 95.23353, which, subtracting the accrued interest of 0.14063 as of this date, equates to a flat price of 95.09290.

Algebraically, then, the full forward price of the 1.125s of 02/15/2031, 95.23353 can be written as,

$$95.23353 = \left[(95.50781 + 0.28280)\left(1 + \frac{0.015\% \times 90}{360}\right) - \frac{1.125}{2}\right] \times \left(1 + \frac{0.015\% \times 46}{360}\right) \tag{11.2}$$

Rearranging terms, and noting that the product of two interest rates is typically very small, gives the following approximation,

$$95.23353 \approx \left[(95.50781 + 0.28280) - \frac{1.125/2}{\left(1 + \frac{0.015\% \times 90}{360}\right)}\right] \times \left(1 + \frac{0.015\% \times 136}{360}\right) \tag{11.3}$$

In words, in the case of an intermediate coupon, the full forward price approximately equals the future value to the forward date of the full spot price less the present value of the coupon payment. More generally, were there more than one intermediate coupon payment, the full forward price would approximately equal the future value of the full spot price minus the present values of all those intermediate coupons.

In both of these examples, forward bond prices are less than spot prices: for the 2.875s of 05/15/2028, the forward price of 109.72 is less than the spot price of 110.77, and for the 1.125s of 02/15/2031, the forward price of 95.09 is less than the spot price of 95.51. This relationship is typical and is commonly known as the *forward drop*.

To understand the intuition behind the forward drop, imagine that a trader has funds equal to the spot price and wants to own the bond as of the forward date. There are two possible strategies, which, by arbitrage, have to be equally appealing: i) use the funds to buy the bond spot and earn coupon interest to the forward date; and ii) buy the bond forward and invest the funds at the repo rate. If the coupon interest from strategy i) exceeds the repo interest from strategy ii), then the two strategies can be equally appealing only if the forward price is less than the spot price. But coupon interest on bonds typically does exceed the repo rate, because the term structure of interest rates is typically upward sloping. Hence, the forward price is usually less than the spot price, and there is usually a forward drop. Of course, if the repo rate exceeds the coupon rate, the argument reverses, and the forward price exceeds the spot price.

Consistent with this intuitive discussion, Appendix A11.1 shows that the forward drop of a bond approximately equals the difference between the interest earned on the bond and the cost of financing its purchase from the settlement date to the forward date. Following the terminology of Chapter 3, then, in which this difference is called *cash carry*, it follows that the forward drop approximately equals cash carry, or, in the jargon of forward and futures markets, simply "carry." This point is illustrated here for these two examples. For the 2.875s of 05/15/2028, the difference between the spot and forward prices is 110.77344 – 109.71721, or about 1.06. At the same time, because the bond earns an actual/actual 2.875% semiannual coupon and finances at an actual/360 repo rate of 0.015%, the difference between the interest earned and the financing cost is about the same: $(2.875/2) \times 136/184 - 0.015 \times 136/360 = 1.06$. For the 1.125s of 02/15/2031, the forward drop is 95.50781 – 95.09290, or 0.41. The difference between the coupon interest and financing cost is about the same: $(1.125/2) \times 136/184 - 0.015 \times 136/360 = 0.41$.

11.2 FORWARD BOND YIELD

As covered in Chapter 3, the yield to maturity of a bond is the single rate such that discounting the bond's cash flows at that rate, from spot settlement to maturity, gives the bond's (full) market price. The forward yield of a bond is defined analogously as the single rate such that discounting the bond's cash flows from forward settlement to maturity by that rate gives the bond's (full) forward price.

Figure 11.4 diagrams the relationship between the spot yield, the repo rate, and the forward yield, and Table 11.3 reports the spot and forward yields of the 2.875s of 05/15/2028 and the 1.125s of 02/15/2031, which are computed using Equation (A3.5) and the data given in the tables of previous section. As explained in that section, when the coupon rate exceeds the repo rate, market prices align such that investors are indifferent between i) buying a bond at its spot price, and ii) investing in repo to the forward date and buying the bond at its lower, forward price. Figure 11.4 and Table 11.3 make the same point in terms of yield: investors are indifferent between i) investing at the spot yield, and ii) investing in repo to the forward date and then at the higher, forward yield. Put another way, the spot yield is a complex weighted average of the relatively low repo rate and the relatively high forward yield, where the weights reflect the lengths of the relative holding periods, as indicated in Figure 11.4.

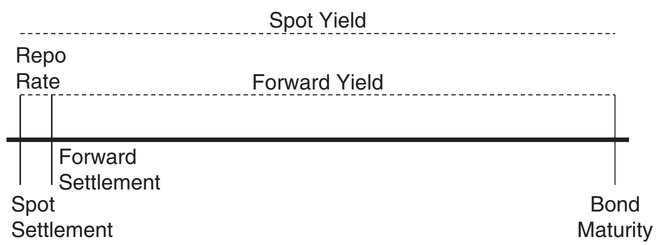

FIGURE 11.4 Spot Yields, Repo Rates, and Forward Yields.

TABLE 11.3 Spot and Forward Yields for the US Treasury 2.875s of 05/15/2028 and 1.125s of 02/15/2031, as of May 14, 2021, for Spot Settlement on May 17, 2021, and for Forward Settlement on September 30, 2021. The Repo Rate is 0.015%. Yields Are in Percent.

Bond	Spot Yield	Forward Yield
2.875s of 05/15/2028	1.260	1.337
1.125s of 02/15/2031	1.625	1.693

11.3 THE INTEREST RATE SENSITIVITY OF A FORWARD CONTRACT

What is the interest rate sensitivity or DV01 of a forward contract? With respect to DV01, and with reference to Figure 11.4, should DV01 be computed by shifting the spot yield, the forward yield, the repo rate, or some combination of the three? Table 11.4 shows the DV01s of the forward contracts in the examples of this chapter with respect to each of these rates. To compute the DV01 with respect to the spot yield, shift the spot yield down one basis point; calculate a new spot price; calculate a new forward price using the new spot price; and take the difference between the new and original forward prices. To compute the DV01 with respect to the forward yield, shift the forward yield down by one basis point; compute a new forward price using the new forward yield; and take the difference between the new

TABLE 11.4 DV01 Metrics for the Forward Prices of the US Treasury 2.875s of 05/15/2028 and 1.125s of 02/15/2031, as of May 14, 2021, for Spot Settlement on May 17, 2021, and Forward Settlement on September 30, 2021. The Repo Rate Is 0.015%.

2.875s of 05/15/2028	
Yield or Rate Shifted	DV01
Spot Yield	0.0708
Forward Yield	0.0666
Repo Rate	–0.0042
1.125s of 02/15/2031	
Yield or Rate Shifted	DV01
Spot Yield	0.0876
Forward Yield	0.0841
Repo Rate	–0.0036

and original forward prices. Finally, to compute the DV01 with respect to the repo rate, shift the repo rate down by one basis point; compute a new forward price using the new repo rate; and take the difference between the new and original forward prices.

The DV01s of the two bonds with respect to the spot yields are consistent with their respective maturities of 7.0 and 9.75 years. The DV01s with respect to the repo rates are negative and small. They are negative, because – from the determination of the forward price in Equation (11.1), (11.2) or (11.3) – the forward price increases as the repo rate increases. And they are small, because the term of the underlying repo between the spot and forward settlement dates is 136 days, or a bit more than 0.37 years. Finally, the DV01s with respect to the forward yields can be understood as follows. A long forward position is equivalent to a long spot position and a short – borrowing cash – repo position. Therefore, the DV01 with respect to the forward yield is the sum of the (positive) DV01 with respect to the spot rate and the (negative) DV01 with respect to the repo rate.

Return now to the question of which of these DV01s should be used, keeping Figure 11.4 in mind. If the spot yield changes while the repo rate is constant, then a shift of the spot yield curve makes sense. If the spot yield and repo rate change together, by the same amount, then a shift of the forward yield makes sense. Taking this thinking a step further, some practitioners – recognizing the empirical regularity that longer-term rates are more volatile than shorter-term repo rates – assume that the repo rate changes by a fixed percentage of the spot yield change. At a percentage of 30%, for example, the DV01 of the forward contract is estimated as 30% times the (negative) DV01 with respect to the repo rate plus the (positive) DV01 with respect to the spot yield, which comes to 0.0695 for the forward on the 2.875s of 05/15/2028 and 0.0865 for the forward on the 1.125s of 02/15/2031. Varying this percentage from 0% to 100%, of course, gives estimates ranging from the DV01s given in Table 11.4 with respect to the spot yields to those with respect to the forward yields, respectively. The best approach, however, is to accept that the repo rate and long-term yields can move independently and, therefore, to hedge in a two-factor framework. Repo rate exposure should be hedged with other short-term, fixed income instruments, and forward yield exposure should be hedged with longer-term instruments.

11.4 MECHANICS OF US TREASURY NOTE AND BOND FUTURES

As mentioned at the start of the chapter, futures and forwards are essentially similar, in that they require the purchase or sale of bonds in the future, but futures embed various *delivery options* and require *daily settlement*.

TABLE 11.5 The Deliverable Basket into TYU1.

Coupon	Maturity	Conversion Factor
2.875	05/15/28	0.8338
2.875	08/15/28	0.8286
1.250	04/30/28	0.7474
1.250	03/31/28	0.7474
3.125	11/15/28	0.8376
2.625	02/15/29	0.8039
2.375	05/15/29	0.7836
1.625	08/15/29	0.7320
1.750	11/15/29	0.7331
1.500	02/15/30	0.7105
0.625	05/15/30	0.6462
0.625	08/15/30	0.6382
0.875	11/15/30	0.6476
1.125	02/15/31	0.6577

To present this material, the text focuses on the 10-year US Treasury note futures contract maturing in September 2021, with the symbol “TYU1.” The “TY” is the code for the 10-year contract; “U” is the code for September; and “1” denotes the last digit of the maturity year, here, 2021. The seller, or short, of one contract commits to sell or deliver $100,000 face amount of one of the *deliverable* Treasury notes listed in Table 11.5 at some time in the *delivery month*, in this case, September 2021. The buyer, or long, commits to buy or take delivery of $100,000 face amount of the Treasury note chosen by the seller at the time chosen by the seller. It is the seller, therefore, that has these two delivery options: the *quality* option, which is the option to choose which note to sell, or deliver, and the *timing* option, which is the option to choose when in the delivery month to sell. There are, in fact, two other delivery options, which are discussed later in the chapter: the *end-of-month* option, which arises because the last trading day of the futures contract (September 21, 2021, for TYU1) is seven business days before the last delivery date (September 30, 2021, for TYU1); and the *wild-card* option, which arises because the futures settlement price on a given day is set hours before an intention to deliver has to be declared.

Once a seller has chosen which bond to deliver and when to deliver it, the price at which the bond is sold is the futures price at the time of delivery times the bond’s *conversion factor*, as listed in Table 11.5. These conversion factors are fixed over the life of the futures contract. The reason for and computation of conversion factors is discussed next. For now, to focus on the mechanics, say that the futures price at expiration is 133.86. If the seller delivers the 2.875s of 05/15/2028, with a conversion factor of 0.8338, the

TABLE 11.6 Settlement Prices of TYU1, from May 10, 2021, to May 28, 2021, and the Daily Settlement to a Long Position of One Contract. Changes Are in Ticks (32nds) and Daily Settlement Payments Are in Dollars.

Date	Price	Change	Daily Settlement
05/10/2021	131–24		
05/11/2021	131–19	−5	−156.25
05/12/2021	131–01	−18	−562.50
05/13/2021	131–10	9	281.25
05/14/2021	131–17+	7.5	234.38
05/17/2021	131–15+	−2	−62.50
05/18/2021	131–16+	1	31.25
05/19/2021	131–04	−12.5	−390.63
05/20/2021	131–19	15	468.75
05/21/2021	131−17+	−1.5	−46.88
05/24/2021	131–23+	6	187.50
05/25/2021	132–02+	11	343.75
05/26/2021	132–01	−1.5	−46.88
05/27/2021	131–25	−8	−250.00
05/28/2021	131–30	5	156.25

price received is $133.86 \times 0.8338 = 111.61$ per 100 face amount. If the seller delivers the 1.125s of 02/15/2031, with a conversion factor of 0.6577, the price received is $133.86 \times 0.6577 = 88.04$. Also, as with any bond sale, the seller through a futures contract receives accrued interest as of delivery.

Futures contracts are subject to daily settlement. Throughout a trading day, market forces determine futures prices, and, at the end of each day, the exchange on which the futures trade determines a *settlement* price. For liquid contracts, like Treasury futures, the settlement price is usually the last traded price of the day. For some contracts, however, in exceptional circumstances, the exchange may substitute its judgment so that the daily settlement price reflects the end-of-day market level. In any case, Table 11.6 reports daily settlement prices for TYU1 from May 10, 2021, to May 28, 2021, per 100 face amount of the underlying bonds. Note that prices are reported in terms of "ticks" or 32nds, so that 131-24 denotes a price of $131 + 24/32$, or 131.75. Note too, that "+" means one half, so that 131-17+ denotes a price of $131 + 17.5/32$, or 131.546875. The third column in the table shows the change in the daily settlement price, also measured in ticks, so that the settlement price fell by five ticks from May 10, 2021, to May 11, 2021.

The daily settlement feature of futures contracts requires that changes in the price of the contract be settled daily. These daily payments, for a long position of one contract in TYU1, are given in the last column of Table 11.6.

From May 10, 2021, to May 11, 2021, for example, because the price of TYU1 falls by five ticks (per 100 face amount), the buyer loses value and must pay $(5/32)\% \times \$100{,}000 = \$156.25$, which the exchange passes on to the seller. From May 12, 2021, to May 13, 2021, on the other hand, the settlement price increases by nine ticks, which means that the seller loses value and must pay $(9/32)\% \times \$100{,}000 = \$281.25$, which the exchange passes on to the buyer. All futures market participants must post *maintenance margin*, usually in the form of a fixed number of dollars per contract, to protect the exchange against their defaulting on daily settlement payments.

With these mechanics explained, the differences between forward and futures contracts can be summarized as follows. First, forward contracts require the purchase and sale of a particular underlying security on a particular date, while futures contracts give the seller the right to choose the security (from a basket of deliverables) and to choose the date (in the delivery month). Second, the profit or loss from a forward contract is realized at the expiration of the contract, while the profit or loss from a futures contract is realized over time. Consider the futures prices in Table 11.6, which, over the period, increase a total of six ticks, from 131-24 to 131-30. Were this a forward contract, the buyer's profit would have increased by six ticks as of the expiration date, in this case, September 30, 2021. But because this is a futures contract with daily settlement, the buyer receives these six ticks as the price changes – some days receiving cash, some days paying cash, but cumulatively, from May 10, 2021, to May 28, 2021, receiving a total of six ticks. A related implication of this reasoning is that, as of May 28, 2021, the forward contract has value – it is a claim on six ticks at expiration – while the futures contract is worth zero – its value has already been fully paid. Put another way, after the settlement payment on May 28, 2021, the buyer of the futures contract on May 10, 2021, at 131-24 has a zero-value position, as does a new buyer of the contract on May 28, 2021, at 131-30.[3]

11.5 PRICING AND HEDGING IMPLICATIONS OF DAILY SETTLEMENT

Daily settlement has implications for the pricing of futures relative to forwards and for hedging using futures rather than forwards. This section considers these two issues in turn.

[3]An historical reason for daily settlement is the mitigation of counterparty risk. Because forwards can accumulate value over time, there can be significant exposure to a counterparty default. With futures, by contrast, exposure is limited to one daily settlement payment, because the contracts of a counterparty that defaults on a settlement payment are canceled. However, with modern practice usually requiring appropriate margin against forward contracts, the two types of contracts no longer differ qualitatively with respect to counterparty risk.

Consider a forward contract and a futures contract on the same, single underlying bond, and assume for the moment that the initial forward and futures prices are the same. From the discussion in the previous section, changes in the prices of the forward contract are realized as profit and loss (P&L) cumulatively, as of expiration, while changes in the prices of the futures contract are realized as P&L daily. Which pattern of P&L realization is preferable to the buyer of a contract?

Because bond prices increase when interest rates fall, and *vice versa*, the buyer of a futures contract realizes profit early when rates fall, that is, when reinvestment opportunities are relatively poor. Similarly, the buyer realizes losses early when rates rise, that is, when the costs of financing those losses are relatively high. In other words, whether rates fall or rise, the early realization of P&L from a long bond futures position is undesirable. Hence, buyers are willing to pay less to purchase bonds through futures than through forwards, and, therefore, the futures price of a bond is less than its forward price. Appendix A11.3 formally proves this result.

In practice, the difference between futures and forward bond prices tends to be small. As evident from the previous paragraph and the appendix to this chapter, the magnitude of the difference depends on the covariance of the underlying bond price with the reinvestment or financing rate to contract expiration.[4] But for note and bond futures, this covariance tends to be low: the time to contract expiration is typically short relative to the maturity of the underlying bonds; short-term rates have relatively low volatility; and short-term rates are imperfectly correlated with long-term rates.

While the futures-forward price difference tends to be small, traders normally do account for the difference in the timing of P&L cash flows when calculating hedge ratios. This is called *tailing* the hedge and is particularly common and useful for *basis trades*, which, as discussed presently, are trades of futures against synthetic bond forwards.

Consider a forward and futures contract on the same underlying security for delivery in d days, when the term repo rate is r. Assume for simplicity that, over a given day, the forward and futures prices increase by the same amount, Δ. The futures contract, then, through daily settlement, pays Δ immediately to the long. The value to the long of the forward contract, by contrast, increases by Δ as of the delivery date, or, in present value terms, by $\Delta/(1 + rd/360)$. Hence, for positive r, less than one futures contract hedges the change in value of one forward contract. More specifically, to hedge N^{fwd}

[4]The covariance of two random variables is the product of their correlation and their two standard deviations. For the covariance to be high, therefore, the variables must be both highly correlated and highly volatile.

forward contracts with N^{fut} futures contracts,

$$N^{fut} = \frac{N^{fwd}}{1 + \frac{rd}{360}} \tag{11.4}$$

where the difference between N^{fwd} and N^{fut} is known as the *tail* of the hedge.

This section closes with a comment on terminology. The terms "mark-to-market," "variation margin," and "daily settlement" are often used interchangeably, although, strictly speaking, they have distinct and different meanings. Mark-to-market is the process of adjusting security prices in an accounting framework to match market values. For example, securities in the trading book of a bank have to be marked-to-market when reported on its balance sheet, while securities designated as "held-to-maturity" can be reported at cost. The term mark-to-market, therefore, does not imply any exchange of cash. Variation margin refers to cash or securities that have to be posted as collateral to secure obligations under a contract. For example, as discussed in Chapter 10, borrowers of cash in the repo market have to make variation margin payments in the form of additional cash or securities as the value of their existing collateral declines. Counterparties posting collateral maintain ownership of that collateral, which means that i) interest is earned on the collateral either by being paid interest on cash collateral or by keeping interest earned on securities posted as collateral, and ii) collateral is returned when the associated contractual obligations have been fulfilled. Daily settlement, as described in this section, refers to the payment of gains or losses. These payments are irrevocable; that is, they do not earn interest and are never returned.[5]

11.6 COST OF DELIVERY AND THE FINAL SETTLEMENT PRICE

The *cost of delivery* measures how much it costs a short to fulfill the commitment to deliver a bond through a futures contract. Having decided to deliver bond i, the short first has to buy the bond at its market price plus accrued interest and then deliver it through the futures contract for the futures price times the conversion factor plus accrued interest. Denoting the time-t flat price of bond i at time t by p_t^i; its accrued interest by AI_t^i; its conversion factor by cf^i; and the futures price by F_t, the cost of delivery is,

$$p_t^i + AI_t^i - (cf^i \times F_t + AI_t^i) = p_t^i - cf^i \times F_t \tag{11.5}$$

[5]The terminology has become even more confusing with interest rate swaps being either collateralized-to-market or settled-to-market. See Chapter 13.

TABLE 11.7 Prices of TYU1 and Notes in Its Deliverable Basket as of the Last Trading Date, September 21, 2021.

TYU1	Price:	133.85938			
Coupon	Maturity	Conversion Factor	Price	Cost of Delivery	Price / Conv. Fac.
2.875	05/15/2028	0.8338	111.61719	0.00524	133.86566
2.875	08/15/2028	0.8286	111.82813	0.91225	134.96032
1.250	04/30/2028	0.7474	101.14063	1.09413	135.32329
1.250	03/31/2028	0.7474	101.17969	1.13319	135.37555
3.125	11/15/2028	0.8376	113.76563	1.64501	135.82333
2.625	02/15/2029	0.8039	110.50000	2.89045	137.45491
2.375	05/15/2029	0.7836	108.82813	3.93592	138.88224
1.625	08/15/2029	0.7320	103.34375	5.35869	141.17999
1.750	11/15/2029	0.7331	104.33594	6.20363	142.32156
1.500	02/15/2030	0.7105	102.20313	7.09604	143.84676
0.625	05/15/2030	0.6462	94.89063	8.39070	146.84405
0.625	08/15/2030	0.6382	94.64844	9.21938	148.30529
0.875	11/15/2030	0.6476	96.57031	9.88298	149.12031
1.125	02/15/2031	0.6577	98.58594	10.54663	149.89499

The short will choose which bond to deliver to minimize the cost of delivery. The bond that minimizes the cost of delivery is called the *cheapest-to-deliver*, or the CTD. Table 11.7 shows the price of TYU1, the prices of the bonds in its deliverable basket, and the costs of delivery for each of those bonds, as of the last trading date, September 21, 2021. For example, the cost of delivery of the 2.875s of 05/15/2028 is,

$$111.61719 - 0.8338 \times 133.85938 = 0.00524 \tag{11.6}$$

Because this is the lowest cost of delivery in the table, these bonds are CTD into TYU1.

To focus on the quality option, assume for the moment that date T is both the last trading date and the last delivery date. Arbitrage arguments can then be used to show that,

$$p_T^{CTD} - cf^{CTD} \times F_T = 0 \tag{11.7}$$

and,

$$F_T = \frac{p_T^{CTD}}{cf^{CTD}} \leq \frac{p_T^i}{cf^i} \tag{11.8}$$

Equation (11.7) says that the cost of delivery of the CTD at expiration is zero. Equation (11.8) says that the futures price at expiration, that is, the last settlement price, equals the ratio of the price of the CTD bond to its conversion factor and that this ratio is less than or equal to the ratios of all other bond prices to their conversion factors. These theoretical predictions are extremely good approximations for TYU1 as of its last trading date, as can be seen from Table 11.7. The cost of delivery of the CTD is very close to zero; the ratio of its price to its conversion factor very nearly equals the futures price; and the ratio of all other bond prices to their CTDs are larger. This section concludes with the arbitrage proofs of Equations (11.7) and (11.8).

Say that (11.7) is not true, and that, instead, $F_T > p^{CTD}/cf^{CTD}$. In that case, an arbitrageur could buy the CTD, sell the futures, and deliver the CTD, earning a profit of,

$$cf^{CTD} \times F_T - p_T^{CTD} \tag{11.9}$$

which is positive by the starting assumption that $F_T > p^{CTD}/cf^{CTD}$. Hence, ruling out market prices that admit riskless arbitrage opportunities rules out this assumed pricing relationship. Next, assume that $F_T < p^{CTD}/cf^{CTD}$. In this case, an arbitrageur could sell the CTD, buy the futures, and take delivery of the bond delivered by the short. If the short delivers the CTD, then the arbitrageur's profit is,

$$p_T^{CTD} - cf^{CTD} \times F_T \tag{11.10}$$

which is positive by the starting assumption that $F_T < p^{CTD}/cf^{CTD}$. If the short delivers some other bond i, the trader buys back the CTD just sold and sells bond i instead, for a total profit of,

$$p_T^i - cf^i \times F_T \geq p_T^{CTD} - cf^{CTD} \times F_T > 0 \tag{11.11}$$

where the first inequality follows from the definition of the CTD – the cost of delivering any bond other than the CTD is at least as great as the cost of delivering the CTD – and the second inequality follows from the starting assumption that $F_T < p^{CTD}/cf^{CTD}$. Hence, the arbitrageur's profit is even greater if the short (suboptimally) delivers a bond other than the CTD. But if the arbitrageur's profit is positive for whatever bond the short delivers, then the assumption $F_T < p^{CTD}/cf^{CTD}$ is ruled out as admitting riskless arbitrage opportunities. Finally, ruling out both $F_T > p^{CTD}/cf^{CTD}$ and $F_T < p^{CTD}/cf^{CTD}$ proves the equality in (11.8) and (11.7). The inequality of (11.8) follows from combining (11.7) with the condition that the CTD has the lowest cost of delivery,

$$
\begin{aligned}
p_T^i - cf^i \times F_T &\geq p_T^{CTD} - cf^{CTD} \times F_T = 0 \\
p_T^i - cf^i \times F_T &\geq 0 \\
F_T &\leq \frac{p_T^i}{cf^i}
\end{aligned}
\tag{11.12}
$$

11.7 MOTIVATIONS FOR A DELIVERY BASKET AND CONVERSION FACTORS

Historically, the design of bond futures contracts purposely avoided a single underlying security. One reason is to ensure that the liquidity of the futures contract does not depend on the liquidity of a single, underlying bond, which might lose its liquidity for idiosyncratic reasons, for example, through being accumulated by a few large, buy-and-hold investors. A second reason is to avoid losing liquidity to the threat of a *squeeze*. A trader squeezes a contract by simultaneously buying many contracts and a large fraction of the deliverable bond, hoping to sell the position at a profit when traders who had sold contracts – but cannot find bonds to deliver – have to pay up to buy the deliverable bond or to buy back the contracts from the perpetrator of the squeeze. The threat of a squeeze, which makes shorts hesitant to take positions, can prevent a contract from attracting volume and liquidity.

A deliverable basket avoids the problems of a single deliverable so long as some bonds can be delivered at a cost not too much greater than the cost of delivering the CTD. In Table 11.7, delivering the 2.875s of 08/15/2028 is about 91 cents per 100 face amount more costly than delivering the CTD, and the 1.250s of 04/30/2028 are about 18 cents more costly than that. The relatively inexpensive substitutability of these near CTD bonds limits the profit potential of a squeeze. After going to all the trouble, expense, and risk of buying many contracts and a large fraction of the outstanding CTD bonds, shorts would refuse to pay a premium of more than about 100 cents per 100 face amount of the CTD, because the two bonds just mentioned can be delivered instead.

Limiting the costs of delivery of at least some bonds other than the CTD is accomplished by conversion factors. In Table 11.7, the costs of delivering the five bonds with the lowest costs of delivery range from about zero to 1.65. This range would be a lot wider, however, if there were no conversion factors, or, more precisely, if all conversion factors were implicitly equal to 1.0. In that case, sellers would receive the futures price for delivering any bond; would choose to deliver the bond with the lowest price; and, therefore, the market futures price would equal that lowest price. Referring back to Table 11.7, the CTD bond would be the 0.625s of 08/15/2030, which has the lowest price of about 94.65, and the futures price would be 94.65. Furthermore, the five lowest bond prices, which range from that 94.65 to the 101.14 price of the 1.25s of 04/30/2028, would imply costs of delivery that range from zero to 101.14 – 94.65 or 6.49, which is a lot wider than the range with the conversion factors of TYU1.

Conversion factors reduce the difference in delivery costs by adjusting delivery prices for the differences across bond coupon rates. Given the limited maturity range of bonds deliverable into TYU1, a major determinant of the price differences across bonds is coupon rate. The most expensive bond in Table 11.7 is the bond with the highest coupon, the 3.125s of 11/15/2028, with a price of 113.77. The least expensive deliverable bond is one of the bonds with the lowest coupon, the 0.625s of 8/15/2030, with a price of 94.65. Conversion factors essentially recognize that sellers are delivering expensive, high-coupon bonds rather than inexpensive, low-coupon bonds. For example, the conversion factor of the 3.125s of 11/15/2028 is 0.8376, while the conversion factor of the 0.625s of 08/15/2030 is 0.6382. A glance at the table reveals that bonds with higher coupons tend to have higher conversion factors.

Conversion factors are computed by the exchange and are easily available. The basic idea can be explained most simply, however, with the following approximation: the conversion factor of a bond equals its price, per dollar face amount, at a yield equal to the *notional coupon rate*, for settlement on the last delivery date. At present, the notional coupon for US Treasury futures is 6%. Therefore, to take one example, the conversion factor of the 1.25s of 03/13/2028 deliverable into TYU1 is approximated using the price-yield Equation (3.8) for a face amount of one dollar, at a yield of 6%, and for settlement on 09/30/2021, which leaves 13 remaining coupon payments,[6]

$$\frac{1.25\%}{6\%}\left(1-\frac{1}{\left(1+\frac{6\%}{2}\right)^{13}}\right)+\frac{1}{\left(1+\frac{6\%}{2}\right)^{13}}=0.7474 \qquad (11.13)$$

This calculation rule clearly assigns higher conversion factors to higher coupon bonds. But the justification for the rule is stronger than that. Assume that the term structure of yields is actually flat at 6%. In that case, the conversion factor of each bond equals the market price of one dollar face amount of the bond, and the ratio of the market price (per 100 face amount) to the

[6]This bond was chosen as the example because it matures in an exact number of semiannual periods. For the other deliverable bonds, use the price-yield Equation (A3.5).

conversion factor equals 100 for every bond. Furthermore, by the arguments of the previous section, the futures price is 100; the cost of delivery of each bond is zero; and all bonds are jointly CTD. In short, if market yields are flat at the notional coupon, conversion factors make every bond as attractive to deliver as any other, thus lessening the dependence of the contract's liquidity on the liquidity of any one bond and also reducing, if not eliminating, the potential profits of a squeeze.

Because the yield curve is not flat at the notional coupon rate, however, conversion factors limit but do not eliminate differences in the costs of delivery across bonds, as evident from Table 11.7. The resulting emergence of a unique CTD is the subject of the next section.

11.8 THE QUALITY OPTION AT EXPIRATION

To illustrate how the quality option works at expiration of the futures contract, assume for the moment that the 2.875s of 05/15/2028 and the 1.125s of 02/15/2013 are the only deliverable bonds into TYU1. Figure 11.5 graphs the ratio of price to conversion factor against yield for these two bonds as of the last trading date, September 21, 2021. The figure assumes that the term structure of yields is flat, that is, that the two bonds have the same yield at the indicated level. At a yield of 6%, the price-conversion factor ratio of each of the bonds is about 100: as explained in the previous section, the conversion factor is approximately equal to the price of the bond (per dollar face amount) at a yield of the notional coupon rate – here 6% – as of the last delivery date (which is only a few days different from the last trading date). Hence, the ratio of each bond's price (per 100 face amount) to its conversion factor is about 100. Furthermore, the futures price is 100, the costs of delivering either bond is zero, and the two bonds are jointly CTD.

As rates move away from 6%, however, the costs of delivering the two bonds diverge. Conversion factors are fixed to equate the price-conversion factor ratio of the two bonds at a yield of 6%. But as yields fall below 6%, the price and price-conversion factor ratio of the shorter-maturity and shorter-duration 2.875s of 05/15/2028 increase more slowly than the price and price-conversion factor ratio of the longer-maturity and longer-duration 1.125s of 02/15/2031. Therefore, for yields below 6%, the 2.875s of 05/15/2028 are CTD. On the other hand, as yields rise above 6%, the price-conversion factor ratio of the longer-duration 1.125s of 02/15/2031 falls more quickly, and that bond is CTD.

The futures price equals the lowest price-conversion factor ratio across deliverable bonds. Graphically, in the present example, the price of TYU1 is the minimum or lower envelope of the two price-conversion factor curves, depicted in Figure 11.5 by a semi-transparent, gray band. At expiration, the

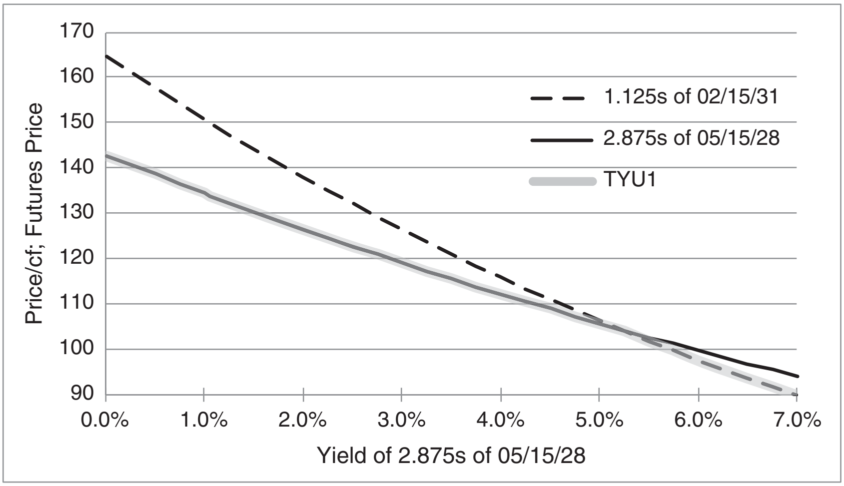

FIGURE 11.5 Quality Option for TYU1 with a Flat Term Structure of Yields, as of the Last Trading Date.

futures price tracks the price of the low-duration bond when rates are low and the price of the high-duration bond when rates are high.

Figure 11.5 is stylized in two ways. First, because there are many deliverable bonds, several distinct ranges of yield at expiration might each determine a different CTD. Along the lines of the discussion here, however, the deliverable bond with the shortest duration is typically CTD at the lowest range of yields; deliverable bonds with intermediate durations might be CTD at intermediate ranges of yield; and the deliverable bond with the longest duration is typically CTD at the highest range of yields.

The second way Figure 11.5 is stylized is in assuming that the yields of the two bonds are equal and move up or down in parallel. In reality, the term structure of interest rates and individual bond yields can move in many ways that affect the determination of the CTD. Most generally, a bond is more likely to be CTD after any change that cheapens it relative to other bonds in the deliverable basket. If the term structure flattens, shorter-term bonds are more likely to be CTD. If the term structure steepens, longer-term bonds are more likely to be CTD. And if the yield of any bond increases idiosyncratically relative to yields of other deliverable bonds, that bond is more likely to be CTD.

11.9 GROSS AND NET BASIS AND BASIS TRADES

The *gross basis* and *net basis* of a deliverable bond measure the difference between the price of that bond and the relevant futures contract. The gross basis of a bond is convenient for quoting the prices of packages, that is, the terms at which traders can buy a futures contract and short a deliverable bond, or *vice versa*. The net basis is a measure of a bond's proximity to being CTD, and changes in a bond's net basis give the P&L of a *basis trade*. Lastly, a basis trade is the purchase of a futures contract and the forward sale of a deliverable bond, or *vice versa*, usually to take advantage of a perceived mispricing of futures relative to the deliverable bonds.

Let p_t^i denote the spot price of bond i at time t; $p_t^i(T)$ its forward price at time t to the last delivery date, T; and cf^i its conversion factor. Let F_t be the futures price at time t. Then, the gross basis and net basis of bond i at time t, GB_t^i and NB_t^i, respectively, are defined as,

$$GB_t^i = p_t^i - cf^i \times F_t \tag{11.14}$$

$$NB_t^i = p_t^i(T) - cf^i \times F_t \tag{11.15}$$

Table 11.8 gives the futures price and the spot and forward prices of each bond deliverable into TYU1, along with its gross basis, carry, and net basis.[7] For example, the gross basis of the 2 7/8s of 05/15/2028 is,

$$\begin{aligned} &(110 - 24\ 3/4) - 0.8338 \times 131 - 17+ \\ &= 110.7734 - 0.8338 \times 131.5469 \\ &= 1.0896 \end{aligned} \tag{11.16}$$

which is 32 times 1.0896, or 34.9 ticks. The net basis of that bond is,

$$\begin{aligned} &109.7172 - 0.8338 \times 131 - 17+ \\ &= 0.0334 \end{aligned} \tag{11.17}$$

or 1.1 ticks.

As defined, the difference between the gross basis and the net basis of a bond equals the difference between its spot and forward prices, which, as shown earlier, equals its carry. This explains the terminology: the net basis is the gross basis net of carry. Continuing with the example, the carry of the 2.875s of 05/15/2028 – from May 17, 2021 (spot settlement), to September 30, 2021 (last delivery date) – is calculated earlier in the chapter as 1.057 or 33.8 ticks. Hence, the bond's gross basis of 34.9 ticks minus its carry of 33.8 ticks equals its net basis of 1.1 ticks. It similarly follows from the definitions that, as of contract expiration – when the forward price for settlement on

[7] As noted already, "carry" in the futures context is "cash carry" as defined in Chapter 3.

TABLE 11.8 Gross and Net Basis of Deliverable Notes into TYU1 as of May 14, 2021. Repo Rate from Spot Settlement to the Last Delivery Date is 0.015%. Gross Basis, Carry, and Net Basis Are in Ticks (32nds).

TYU1	**Price:**	**131-17+**					
Coupon	**Maturity**	**Conv. Factor**	**Spot Price**	**Gross Basis**	**Fwd Price**	**Carry**	**Net Basis**
2.875	05/15/2028	0.8338	110-24 3/4	34.9	109.7172	33.8	1.1
2.875	08/15/2028	0.8286	110-27	59.0	109.7759	34.2	24.8
1.250	04/30/2028	0.7474	99-26 1/4	48.1	99.3615	14.7	33.4
1.250	03/31/2028	0.7474	99-29 1/4	51.1	99.4578	14.6	36.5
3.125	11/15/2028	0.8376	112-22 1/4	80.4	111.5468	36.8	43.6
2.625	02/15/2029	0.8039	109-03+	107.5	108.1348	31.2	76.3
2.375	05/15/2029	0.7836	107-07 1/4	132.7	106.3549	27.9	104.8
1.625	08/15/2029	0.7320	101-13+	164.1	100.8205	19.2	144.9
1.750	11/15/2029	0.7331	102-07+	185.5	101.5934	20.5	165.0
1.500	02/15/2030	0.7105	99-27 1/4	204.4	99.2968	17.8	186.6
0.625	05/15/2030	0.6462	92-05 1/4	229.1	91.9383	7.2	221.8
0.625	08/15/2030	0.6382	91-24 3/4	250.2	91.5451	7.3	242.9
0.875	11/15/2030	0.6476	93-20 1/4	270.2	93.3147	10.2	260.0
1.125	02/15/2031	0.6577	95-16 1/4	287.7	95.0929	13.3	274.4

contract expiration is the same as its spot price at contract expiration – gross and net basis are equal to each other and to the bond's cost of delivery, as defined in Equation (11.5). Furthermore, because the cost of delivery at expiration of the CTD is zero, its gross and net basis at expiration are also zero.

Gross and net basis are particularly useful in the context of basis trades. To buy or be long a bond's basis, a trader buys a bond forward to contract expiration and sells a *conversion-factor weighted* number of futures contracts. To sell or be short a bond's basis, a trader sells a bond forward to contract expiration and buys a conversion-factor weighted number of futures contracts. In the US Treasury market, where spot and repo markets are much more liquid than forward markets, forward purchases and sales of bonds are typically synthetic, as described earlier in the chapter. A conversion-factor weighted number of contracts is the number of contracts corresponding to the face amount of the bonds times the bond's conversion factor. For example, a trader buying $100 million of the 2.875s of 05/15/2028 TYU1 basis buys $100 million face amount of the bonds forward and sells 0.8338 times $100 million divided by $100,000, or about 834 contracts.

Gross basis is often used to quote a package trade of bonds against futures. A trader might put in an order, for example, to buy $100 million face amount of the 2.875s of 08/15/2028 and sell 834 TYU1 contracts at a gross basis of 34.9 ticks or less. If prices subsequently line up such that the bond price minus 0.8338 times the futures price is less than or equal to 34.9 ticks, the trade is executed. The trader then sells the repo to TYU1's expiration date to complete the basis trade.

Abstracting for a moment from interest on daily settlement payments, the P&L of a long basis position is the sum of the profit or loss from the forward bond position and from the futures position. Algebraically, with G^i face amount of bond i and a trade from time t to time s, the P&L is,

$$G^i \times [P_s^i(T) - P_t^i(T)] - G^i \times cf^i \times [F_s - F_t] \tag{11.18}$$

But, from the definition of the net basis in Equation (11.15), (11.18) can be rewritten as,

$$G^i \times [NB_s^i(T) - NB_t^i(T)] \tag{11.19}$$

In words, the P&L of a long basis trade is the face amount of bonds times the change in the bond's net basis. The P&L of a short basis trade is the negative of that. Consider, then, a trader who is sure that a particular bond will be CTD at expiration but sees that the net basis of the bond is trading at five ticks. This trader might very well sell the basis in the hope of making five ticks: if the bond does turn out to be CTD, its net basis will be zero at expiration, and the trader's profit will be the difference between the initial net basis of five ticks and the final net basis of zero ticks. Along these lines, therefore, the net basis of a bond is an indicator of a bond's becoming CTD. With respect to TYU1 as of May 14, 2021, the net bases in Table 11.8 strongly indicate that the 2.875s of 05/15/2028 will be CTD at expiration. Price are such that traders cannot make much money by betting that the 2.875s of 05/15/2028 will be CTD, that is, by selling its basis at only 1.1 ticks. But they can make a lot of money by taking a contrarian view that turns out to be right, for example, by selling the 2.875s of 08/15/2028 basis at 24.8 ticks in the hope that this bond will turn out to be CTD.

The interpretation of net basis in the previous paragraph implicitly assumes that the futures price is fair, or valued correctly, relative to the prices of the underlying bonds and their volatility. In terms of TYU1 as of May 14, 2021, assuming that the futures price is fair is equivalent to saying that the net basis of the 2.875s of 05/15/2028 is 1.1 ticks – rather than zero – because there is still some chance that another bond will turn out to be CTD at delivery. But some trader might believe that bond price volatility from May 14, 2021, to September 30, 2021, is much too small for a change in the CTD to be possible. From that trader's perspective, the market's net basis of the 2.875s of 05/15/2028, at 1.1 ticks, is too high; that is, the price of the 2.875s of 05/15/2028 is too high relative to the futures price. In this sense, net basis can be an indicator of relative value.

This section closes with two details about gross and net basis and basis trades. First, Equations (11.18) and (11.19) are not completely accurate expressions of the P&L of basis trades as described, because interest is paid or earned on daily settlements of the futures contract. As mentioned earlier, however, tailing the hedge does account for daily settlements. Therefore, these P&L equations are accurate representations of P&L for basis trades with a tail hedge, that is, not with a conversion-factor weighted number of futures contracts, but with a conversion-factor weighted and tailed number of futures contracts. Second, some traders, instead of selling or buying the repo to the term of the futures contract, sell or buy overnight repo. These traders are not really buying or selling the bonds forward against their futures positions, and the resulting P&L is not equal to that given in Equations (11.18) and (11.19). The case study at the end of this chapter discusses the risks of a "basis" trade with overnight rather than term repo.

11.10 IMPLIED REPO RATES

Table 11.1 and Equation (11.1) compute the forward price of the 2.875s of 05/15/2028 given its price for spot settlement and its term repo rate from spot to forward settlement. The same relative pricing formula can be applied, of course, to compute the repo rate given a bond's forward and spot prices. Under the assumption that a particular bond will be CTD into a futures contract, a repo rate computed in this way is called that bond's *implied repo rate*.

To illustrate, consider, once again, the 2.875s of 05/15/2028 and the data in Table 11.8. If that bond will be CTD with certainty, the ratio of its forward price to its conversion factor should equal the futures price; that is, its forward price should equal 131 – 17 + /0.8338, or 109.68379.[8] Then, following the logic of Equation (11.1), the relationship between the bond's spot and forward prices and the repo rate, r, is,

$$109.68379 + 1.07813 = (110.77344 + 0.01562)\left(1 + \frac{r \times 136}{360}\right) \quad (11.20)$$

which gives an implied repo rate of $r = -0.065\%$.

It is evident from Equations (11.1) and (11.20) that a lower repo rate generates a lower forward price. Therefore, if a bond will be CTD with certainty, and if its implied repo rate is less than its actual repo rate, then the futures price is cheap relative to the bond's spot price. Similarly, if a bond will be CTD with certainty, and if the implied repo rate of a bond is greater than its actual repo rate, then the futures price is rich relative to the

[8]The definition of implied repo ignores the futures-forward price difference.

bond's spot price. In the example, under the assumption that the 2.875s of 05/15/2028 will be CTD, an implied repo rate of −0.065% and an actual repo rate of 0.015% imply that TYU1 is cheap relative to the CTD's spot price.

Another way to think about implied repo is as the rate earned by buying a bond spot and selling it through the futures contract. Rewriting Equation (11.1), with a spot price of 110.77344 and a forward price of 109.71721, the return from buying spot and selling forward is just the repo rate,

$$\frac{109.71721 + 1.07813 - (110.77344 + 0.01562)}{(110.77344 + 0.01562)} \frac{360}{136} = 0.015\% \quad (11.21)$$

Similarly, the return from buying spot and selling forward – this time with a forward price implied by the futures price of 109.68379 – is −0.065%. From this perspective as well, then, TYU1 is too low relative to the spot price of the 2.875s of 05/15/2028, so long as that bond is CTD. Implied repo as a measure of relative value will be revisited in the case study at the end of the chapter.

11.11 FUTURES PRICE AND THE QUALITY OPTION BEFORE EXPIRATION

It is shown earlier, and illustrated in Table 11.7, that the futures price at expiration is the minimum ratio of price to conversion factor across deliverable bonds. This section analyzes the futures price before expiration. The basic idea is as follows. If the CTD were known with certainty, then (abstracting from the futures-forward difference) the futures price would equal the forward price of the CTD divided by its conversion factor. But because the CTD is not known with certainty, and because the seller of the futures contract has the option to choose which bond to deliver, the futures price is reduced by the value of that delivery option.

To illustrate this idea graphically, Figure 11.6 shows a yield-price curve for TYU1, as of May 14, 2021, along with the ratios of forward prices to conversion factors for two deliverable bonds – the shortest-duration bond in the basket, the 2.875s of 05/15/2028, and the longest-duration bond in the basket, the 1.125s of 02/15/2031. The horizontal axis gives the assumed yield of the 2.875s of 05/15/2028 as of May 14, 2021, which for graphical clarity, is limited to the range of 4.00% to 5.50%. The yield of the 1.125s of 02/15/2031 and the repo rate are assumed to be at fixed spreads to the yield of the 2.875s of 05/15/2028.[9] The futures price corresponding to each

[9]On May 14, 2021, the yield of the 2.875s of 05/15/2028 was 1.26%; the yield of the 1.125s of 02/15/2031 was 36.5 basis points higher, at 1.625%; and the repo rate

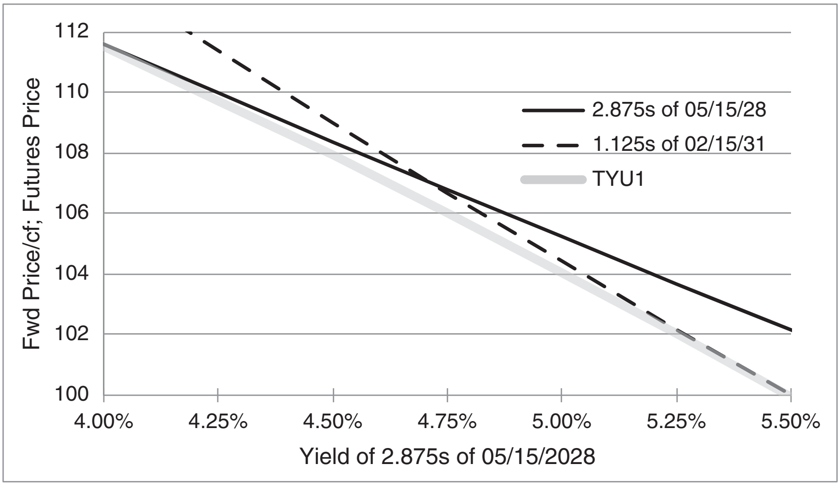

FIGURE 11.6 TYU1 Price with Two Deliverable Bonds, as of May 14, 2021. Yields Are Assumed to Move in Parallel with a Volatility of 100 Basis Points per Year.

yield of the 2.875s of 05/15/2028 is calculated using a very simple model, which assumes that i) changes in the yield of the 2.875s of 05/15/2028 from its starting value along the horizontal axis to the expiration date, September 30, 2021, are normally distributed with a volatility of 100 basis points per year; ii) spreads of the yield of the 1.125s of 02/15/2031 and the repo rate to the yield of the 2.875s of 05/15/2028 are fixed; and iii) delivery occurs on September 30, 2021; that is, only the quality option is considered. The appendix to this chapter describes the pricing of futures in a term structure model in more detail.

There are only four and a half months from the pricing date of the figure, May 14, 2021, to the last delivery date, September 30, 2021. Therefore, when yields are very low on May 14, 2021, it is extremely likely that the relatively low-duration 2.875s of 05/15/2028 will be CTD at expiration. Or, put differently, it is extremely unlikely that rates will rise enough for the relatively high-duration 1.125s of 02/15/2031 to become CTD. Recognizing this, the market assumes that the 2.875s of 05/15/2028 will be delivered and sets that bond's net basis very close to zero, or, equivalently, sets the futures price very near the ratio of that bond's forward price to its conversion factor. This can be seen in Figure 11.6: for low rates, the semi-transparent, gray band representing the futures price converges to the forward price-conversion factor ratio of the 2.875s of 05/15/2028. Analogously, when rates are very high on

---

was 124.5 basis points below, at 0.015%. These are the spreads kept fixed in the figure.

May 14, 2021, it is extremely likely that the relatively high-duration 1.125s of 02/15/2031 will be CTD; its net basis is near zero; and the futures price converges to the ratio of its forward price to its conversion factor.

For intermediate rates shown in the figure, the futures price is less than the forward price-conversion factor ratio of both bonds. Buyers will not price the futures under the assumption that the 2.875s of 05/15/2028 are delivered with certainty, because the sellers – who minimize the value of the futures – will switch from delivering the 2.875s of 05/15/2028 to delivering the 1.125s of 02/15/2031 exactly in those scenarios that hurt the buyers. Hence, the futures price must be lower than it would be under the assumption that the 2.875s of 05/15/2028 will be delivered with certainty. By the same logic, the futures price must be lower than it would be under the assumption that the 1.125s of 02/15/2031 will be delivered with certainty. In fact, the distance between either bond curve and the futures price curve at a given yield represents the value of the seller's option to switch from delivering one bond to delivering the other. That value is low when rates are very high or very low, because one bond or the other is almost certain to be delivered. Furthermore, if the assumed volatility of yields were higher, so that yields are more likely to move far from their levels as of the pricing date, the option to switch would increase in value, and the futures curve would be even more below the bond curves.

Figure 11.7 illustrates the value of the quality option as of May 14, 2021, for yields as of that pricing date ranging from 0% to 7%, in terms of the net basis of three bonds: the low- and high-duration bonds described earlier, and an intermediate-duration bond in the basket, the 3.125s of 11/15/2028. These curves are computed using the same model as in Figure 11.6.

The net basis of the low-duration bond, the 2.875s of 05/15/2028, is near zero for low yields, because at those levels of yields on May 14, 2021, it will almost certainly be the CTD at expiration. As rates increase, however, because it becomes more likely that the 1.125s of 02/15/2031 will be CTD, the net basis of the 2.875s of 05/15/2028 increases. Put another way, this net basis increases because the value of the seller's option to switch away from delivering the 2.875s of 05/15/2028 increases. In contrast, the net basis of the high-duration bond, the 1.125s of 02/15/2031, is near zero for high yields, when it will almost certainly be CTD at expiration. As yields fall, however, its becoming CTD is less of a certainty, and its net basis increases.

The net basis of the 3.125s of 11/15/2028 is never close to zero, because it is never likely to become CTD. While the set of deliverables can, in general, lead to ranges of yields in which intermediate-duration bonds are CTD at expiration, for TYU1 as of May 14, 2021, and under the assumption of parallel shifts, the CTD at expiration is heavily likely to be either the lowest-duration bond, the 2.875s of 05/15/2028, or the highest-duration bond, the 1.125s of 02/15/2031. Nevertheless, at a yield level of about 4% as of the pricing date, the 3.125s of 11/15/2028 are as close as they ever are

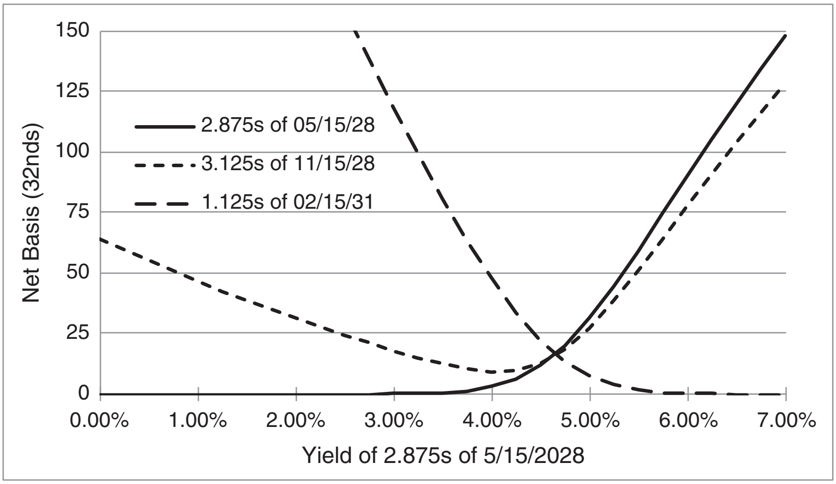

FIGURE 11.7 Net Basis of Three Bonds Deliverable into TYU1, as of May 14, 2021. Yields Are Assumed to Move in Parallel with a Volatility of 100 Basis Points per Year.

to being CTD at expiration, and their net basis falls to a minimum of about nine ticks. For lower yields and for higher yields, the 3.125s of 11/15/2028 are even less likely to wind up as the CTD, as their net basis increases accordingly.

The net basis graphs in Figure 11.7, as manifestations of the quality option, can be described more directly in terms of option terminology. The net basis of the low-duration bond resembles a call option on rates, or, equivalently, a put option on bond prices; the net basis of the intermediate-duration bond resembles a straddle on rates and bond prices; and the net basis of the high-duration bond resembles a put on rates or a call option on bond prices.

The price-yield curve of TYU1 in Figure 11.6, which incorporates the seller's quality option, can be used to derive the DV01 of the futures contract, as defined in Chapter 4: the change in price for a one-basis-point decline in rates. The DV01s of the 2.875s of 05/15/2028 and 1.125s of 02/15/2031, divided by their respective conversion factors, along with the DV01 of TYU1, are graphed Figure 11.8. As expected, the DV01 of the shorter-maturity bond is relatively low and flat, with the flatness indicative of its lower convexity, while the DV01 of the longer-maturity bond is relatively high and steep, with the steepness indicative of its higher convexity. The futures price-yield relationship, however, differs markedly from that of a coupon bond. At low yields as of the pricing date, the futures DV01 resembles the DV01 divided by conversion factor of the bond most likely to be CTD at expiration, namely the shorter-maturity 2.875s of 05/15/2028. Analogously, at high yields as of the pricing date, the futures DV01 resembles that

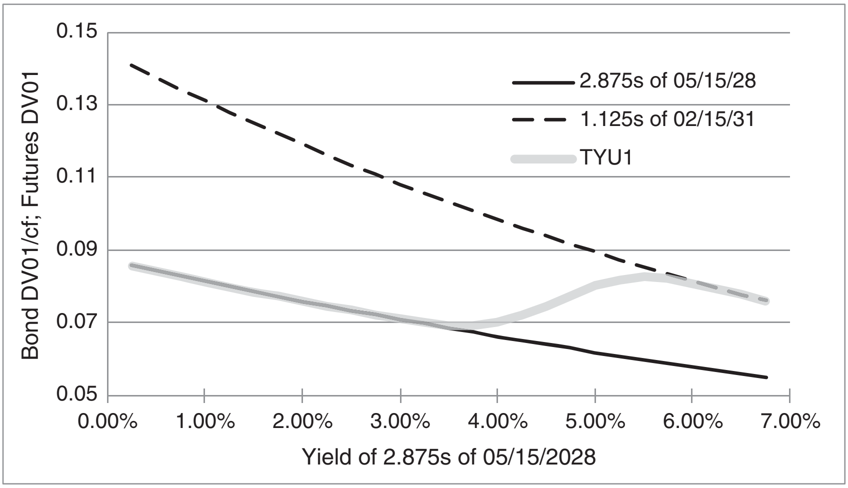

FIGURE 11.8 DV01 of TYU1 and the DV01s of Two Deliverable Bonds Divided by Their Conversion Factors, as of May 14, 2021. Yields Are Assumed to Move in Parallel with a Volatility of 100 Basis Points per Year.

of the 1.125s of 02/15/2031 divided by its conversion factor. At intermediate yields, the DV01 of the futures contract gradually changes from following that of the short-maturity bond to following that of the high-maturity bond. Therefore, its DV01 increases as yields increase, which means that TYU1 is negatively convex over that region of intermediate yields. This negative convexity is also visible, of course, in the shape of the TYU1's price-yield curve in Figure 11.6.

As of May 14, 2021, the yield of the 2.875s of 05/15/2028 was, in fact, 1.26%. In this low-rate environment, it is clear from Figures 11.6, 11.7, and 11.8 that the 2.875s of 05/15/2028 are solidly CTD: the quality option is worth very little; the futures contract – apart from daily settlement considerations – is essentially a forward contract on the 2.875s of 05/15/2028; and the futures price is approximately equal to the forward price of that bond divided by its conversion factor. These conclusions are all evident in Table 11.8, which shows that, as of May 14, 2021, the net basis of the 2.875s of 05/15/2028 is 1.1 ticks, and, therefore, the futures price of 131-17+, or about 131.55, is very close to the ratio of that bond's forward price to its conversion factor, 109.7172/0.8338, or 131.59.

These observations raise the question of why the notional coupon of the contract has not been lowered from 6%. As discussed already, when the yield curve is flat at 6%, conversion factors make all bonds equally deliverable, which, in turn, lowers liquidity dependence on a single CTD and reduces the profitability of a squeeze. That the term structure is not flat, and that bonds sell at idiosyncratically variable yields means that conversion factors are not perfect with respect to these objectives, but setting the notional coupon close to prevailing market levels does significantly reduce differences in deliverability across bonds. Again, then, why has the notional coupon not been set closer to market levels, that is, at somewhere between 1% and 2%? After all, the current notional coupon of 6% was reduced from its previous level of 8% in response to lower market rates.

A likely answer has two parts. First, with the increase of volumes and liquidity of both Treasury bonds and Treasury futures, the risks of a fall in liquidity or a squeeze have diminished. Second, as evident from the discussions in this chapter, valuing the quality option is quite complex. Overall, market participants – apart from basis arbitrage traders – seem to prefer a simpler contract, one effectively without delivery options, which is, therefore, easier to price and hedge.

11.12 THE TIMING, END-OF-MONTH, AND WILD-CARD OPTIONS

As mentioned earlier, the timing option allows the seller to deliver at any time during the delivery month. For TYU1, for example, delivery is allowed at any date between September 1, 2021, and September 30, 2021. To understand whether sellers should deliver late or early, consider a trader who is long the bond and short the futures. Delivering late earns carry on the bond coupon until delivery and preserves full use of the quality option. Delivering early and investing the proceeds at the repo rate gives up carry on the bond and any remaining value of the quality option. Therefore, so long as carry is positive, it is optimal to delay delivery. Because coupon rates are typically greater than repo rates, carry is typically positive and late delivery typically optimal. If carry is negative, however, which is most likely for shorter-maturity contracts when the short end of the term structure is downward sloping, then the carry advantage of early delivery has to be weighed against the lost option value.

Before analyzing the end-of-month option, the discussion turns to the fact that the last trading date of Treasury note and bond futures is seven business days before the last delivery date. For TYU1, for example, the last trading date is September 21, 2021, while the last delivery date is September 30, 2021. This feature of the contract requires an adjustment to futures hedges just after the last trading date. To see this, consider a basis position on the last trading date. For simplicity, continuing with TYU1, assume that the 2.875 of 05/15/2028 is almost certain to be CTD, and that, with the delivery date so near, its forward and spot prices are approximately equal. Now consider a $1 billion long basis position in that CTD, which is long $1 billion face amount of the bond and short 8,338

TYU1 contracts. This position is hedged because the futures price equals the CTD bond price divided by its conversion factor, 0.8338. If, for example, the price of the bond falls from 111.62 to 110.62, causing a loss on the (long) bond side of the trade of \$1 billion × (111.62% − 110.62%), or \$10 million, the futures price falls 8,338 × \$100,000 × (111.62%/0.8338 − 110.62%/0.8338), for a gain on the (short) futures side of the trade of an offsetting \$10 million. After the last trading date, however, the futures price is frozen at its last settlement price, say of 111.62/0.8338, or 133.8690, and the delivery price of the CTD is frozen at 133.8690 × 0.8338 = 111.62. If the CTD now falls in price to 110.62, the value of the long bond position again falls by \$10 million, but the short futures position gains nothing. Put another way, because the basis trader is short 8,338 contracts, only 8,338 times \$100,000, or \$833.8 million face amount of the CTD can be sold at the fixed futures settlement price of 111.62. The remaining \$166.2 million long face amount of the CTD suffers a loss from its fall in price of \$166,200,000 × (111.62% − 110.62%), or \$1.662 million.

Given the impact of freezing the futures price as of the last trading date, the futures hedge has to be changed from a conversion-factor weighted number of contracts to a matching notional amount of contracts. In the example, after the last daily settlement on September 21, 2021, at a futures price of 133.8690, the long basis trader can sell an additional 1,662 TYU1 contracts for a total short of 1,000 contracts. From then on, the position is hedged because the trader can sell all \$1 billion face amount of the bonds through the futures contract at a fixed price of 133.8690 × 0.8338 = 111.62.[10]

Returning now to the end-of-month option, say that, after the last trading date, the futures price is frozen at F_s and that a long basis trader is long bond i and short a matching notional amount of contracts. This trader, planning to deliver bond i, has locked in a delivery price of cf^i times F_s. But a profit can be earned by switching to bond j, that is, by selling bond i, buying bond j, and delivering bond j, if there is a bond j such that,

$$p^i - p^j + cf^j \times F_s > cf^i \times F_s \tag{11.22}$$

$$p^i - cf^i \times F_s > p^j - cf^j \times F_s \tag{11.23}$$

Equation (11.22), which says that the switch is profitable, is equivalent to Equation (11.23), which says that bond j is cheaper to deliver than bond i.

[10]To avoid confusion in terminology, note that many practitioners refer to the difference between the matching notional amount and the conversion-factor weighted number of contracts as a "tail," but this hedging adjustment is not at all related to "tailing the hedge" to account for the timing of daily settlement payments.

The end-of-month option – switching from bond *i* to bond *j* in the previous paragraph – is potentially valuable, but it often turns out not to be worth much in practice. First, since the time between the last trading and last delivery dates is relatively short, relative bond prices do not tend to change very much. Second, traders who are long the basis actively monitor opportunities to profit by switching into bonds that are cheaper to deliver. As a result, any time a deliverable bond does cheapen relative to others, many traders have an incentive to buy that bond to switch and the cheapening comes to an abrupt halt.

The last of the delivery options, the *wild-card* option, arises because futures prices settle every day at 2 p.m. CT, but notice of delivery is due a few hours later. Consider a long $1 billion CTD basis (i.e., the 2.875s of 05/15/2028) into TYU1 sometime during the delivery month, but before the last trading date. As discussed before, before the last trading date, the long CTD position of the basis trade is hedged against changes in its price by a short of 8,338 contracts. Now, if on a particular day the price of the CTD rises significantly between the time the futures price settles and the time notice of delivery is due, that is, if the bond price rises while the futures settlement price from 2pm is still applicable, the trader can profit by selling $166.2 million face amount of the CTD at its now higher price and giving notice to deliver the remaining $833.8 million face amount through the futures at the now stale settlement price. The profit from this liquidation of the position is the increase in the value of the $166.2 million CTD face amount. Whether exercising the wild-card option is worthwhile depends on how much the price of the CTD increases relative to how much carry is sacrificed by exercising the timing option early.

11.13 CASE STUDY: BASIS TRADES IN MARCH 2020

In the wake of the COVID pandemic and economic shutdown, US Treasury yields fell during the second half of February and early March 2020, in a relatively orderly flight to safety. Starting March 9, however, many market participants sold Treasury bonds to raise cash, temporarily reversing the decline in yields and ushering in a period of heightened market volatility. These market movements are illustrated in Figure 11.9, which shows yields of the CTD bonds into the June 2020 two-year (TUM0), five-year (FVM0), and 10-year (TYM0) Treasury note futures contracts.

For a couple of years before the pandemic, Treasury futures had often traded somewhat rich to their underlying bonds. Many attributed this relative mispricing to a preference for futures over bonds by regulated financial institutions seeking to limit the sizes of their balance sheets: long futures positions do not appear on the balance sheet, while purchased bonds do. In

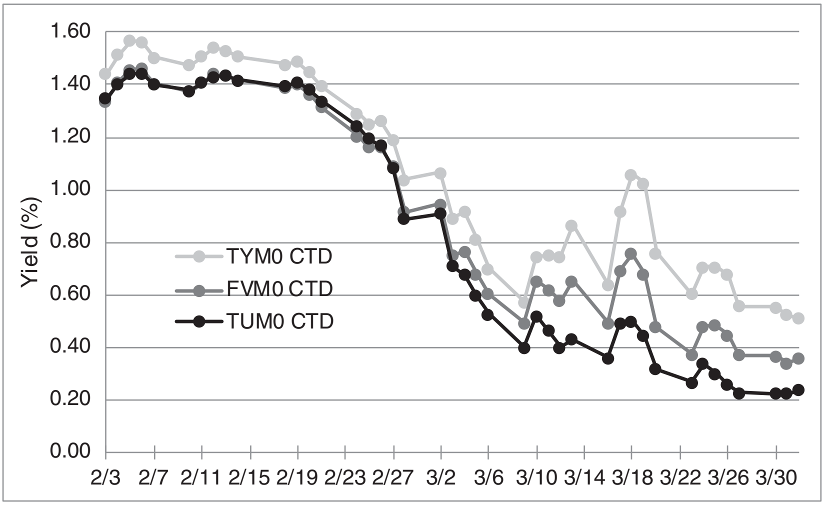

FIGURE 11.9 Yields of CTD Bonds into TYM0, FVM0, and TUM0, from February 3, 2020, to April 1, 2020.

any case, Figure 11.10 shows that, through much of February 2020, TYM0 traded a bit rich to its CTD, in that the net basis of the CTD into TYM0 fluctuated between zero and minus five ticks. Figure 11.11 shows that TUM0 and FVM0, along with TYM0, traded slightly rich in that the implied repo rate of the CTD bonds into these contracts exceeded the actual repo rate, represented by SOFR.[11]

Responding to the chronic richness of futures contracts, hedge funds and their relative value traders consistently bought the CTD basis. To understand the attractiveness of this trade, consider a hedge fund that on February 10, 2020, buys $1 billion face amount of the TYM0 CTD basis at −4.0 ticks, which is consistent with the levels shown in Figure 11.10. Given that interest rates are almost certainly not going to rise enough to change the CTD, the net basis will be zero at expiration, and the profit on the trade would be (4/32)% times $1 billion, or $1.25 million. If levered 20-to-1, which many believe was typical for trades of this type, the hedge fund would assign $50 million of capital to the trade and earn a return on capital, from February 10 to June 30 (141 days), of $1.25 million/$50 million or 2.5%, which is about 6.5% annualized. If levered 40-1, which has been characterized as common

[11]Note that the net basis in Figure 11.10 and the implied repo rate in Figure 11.11 assume that the term repo rate on each day equals the overnight repo that day. This is not a great assumption but is necessitated in this case by the lack of data on term repo rates.

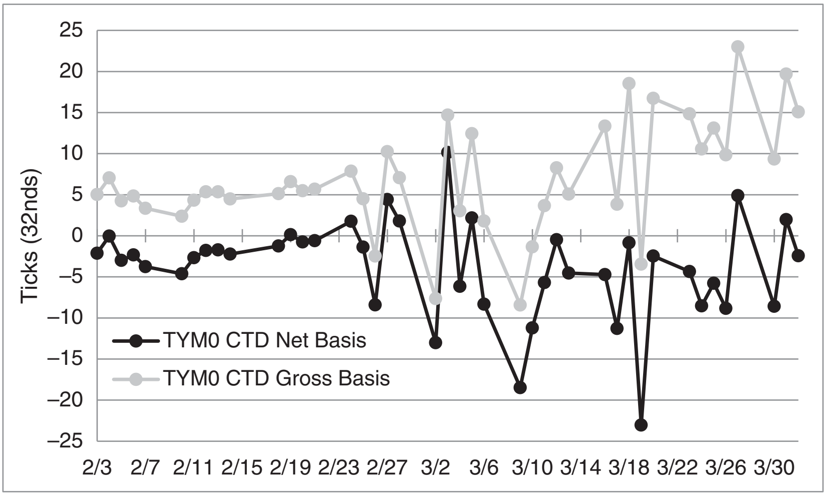

FIGURE 11.10 Net and Gross Basis of CTD into TYM0, from February 3, 2020, to April 1, 2020.

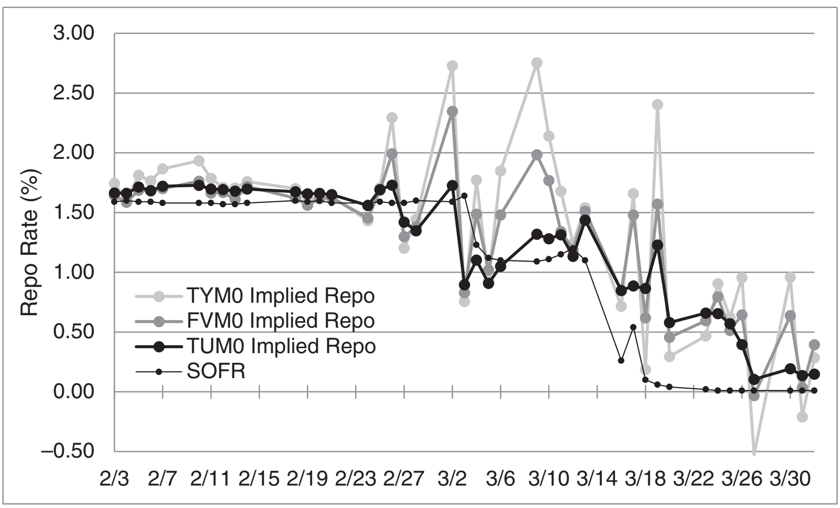

FIGURE 11.11 Implied Repo Rates of CTD Bonds into TYM0, FVM0, TUM0, and SOFR from February 3, 2020, to April 1, 2020.

for these trades, assigned capital would be $25 million and the annualized return would be about 13%.[12]

Buying the basis at a negative price and holding the position until the net basis converges to zero had indeed, for some time, been a relatively reliable and stable money machine. From March 9, however, when Treasury yields became much more volatile, so did the P&L of long basis trades. The newfound volatility of the trade can be seen from the net and gross basis of TYM0 in Figure 11.10 and from the implied repo of TYM0, FVM0, and TUM0 in Figure 11.11. The market dislocations at that time can also be seen from rising spreads between CTD yields and the yields of nondeliverable, but otherwise comparable bonds. The TUM0 line in Figure 11.12 shows the difference between the yields of the 1.875s of 02/28/2022 and the CTD into TUM0, that is, the 2.375s of 03/15/2022. While these two bonds traded at a low and stable spread through February and early March, that spread rose and became much more volatile over the rest of March. Similarly behaved are the FVM0 and TYM0 spreads, that is, the spread between the 1.875s of 08/31/2024 and the CTD into FVM0, that is, the 1.25s of 08/31/2024, and the spread between the 2s of 11/15/2026 and the CTD into TYM0, that

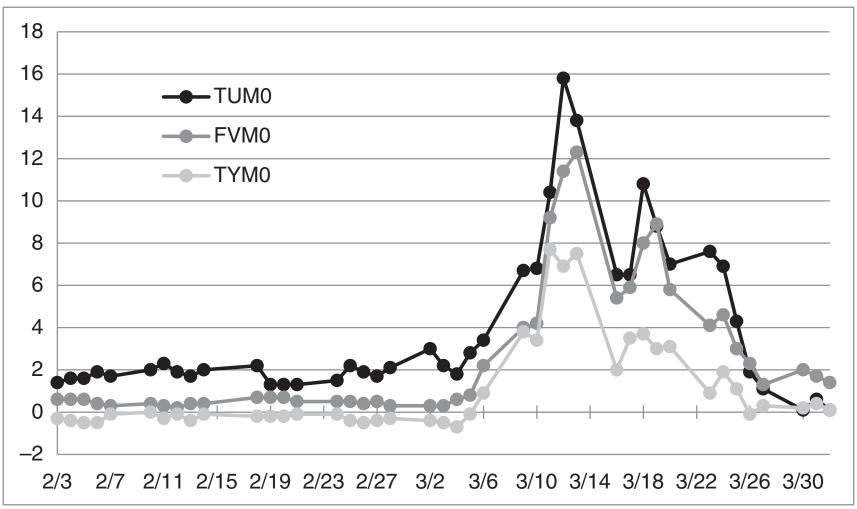

FIGURE 11.12 Richness of CTD Bonds into TYM0, FVM0, and TUM0 Relative to Comparable, Non-Deliverable Bonds, from February 3, 2020, to April 1, 2020.

[12]Assigned capital in this context is an allocation from a fund's scarce risk and funding resources to one of its various trading or investment strategies. As the following discussion shows, capital assigned in this way must be available in ready funds should the need arise.

TABLE 11.9 A $1 Billion Long TYM0 CTD Basis Trade, from March 6, 2020, to March 12, 2020. Net Basis Is in Ticks (32nds). All Other Columns Are in Dollars.

Date	Net Basis	Daily P&L	Cumulative P&L	Margin per Contract	Margin Call	Cumulative Cash
03/06/2020	–8.33			1,275		–10,171,597
03/09/2020	–18.48	–3,172,148	–3,172,148	1,375	–797,772	–14,141,517
03/10/2020	–11.22	2,268,613	–903,535	1,375	0	–11,872,904
03/11/2020	–5.70	1,725,636	822,101	1,600	–1,794,988	–11,942,256
03/12/2020	–0.47	1,633,228	2,455,330	1,600	0	–10,309,028

is, the 2.25s of 02/15/2027. Generally speaking, liquidity is in high demand when markets are volatile, and the most liquid bonds tend to trade at a premium. CTD bonds, which inherit the liquidity of their associated futures contracts, trade in volatile markets at a premium relative to bonds that – in normal times – are comparable. This premium, however, like the demand for liquidity, can be quite volatile, as evident from Figure 11.12.

Long basis trades seem riskless, so long as they can be held to expiration. Through volatile markets, however, traders might not actually be able to maintain their positions. To illustrate, Table 11.9 describes a $1 billion long TYM0 CTD basis trade initiated at −8.33 ticks on March 6, 2020, and held through March 12, 2020. The table assumes that a trader buys $1 billion face amount of the bonds; sells a conversion-factor weighted and tailed number of futures contracts, in this case, 7,978 contracts;[13] and sells the CTD repo to the expiration of the contract. For this trade, therefore, as explained earlier in the chapter, the daily P&L, measured as of the delivery date, equals the change in the net basis times the $1 billion face amount of the bond position.

The net basis closed at about −18.5 ticks on March 9, generating a loss of about $3.2 million. This is a sizable loss given that the target horizon profit of the trade is only the original 8.33 ticks, that is, (8/32)% times $1 billion, or $2.5 million. Many hedge funds manage risk, at least in part, by loss limits, and a loss of this magnitude might result in a trader being forced to exit the position. From then on, Table 11.9 reports that the trade makes money, as the net basis rises to nearly zero. By reporting only daily closes, however, the table minimizes the true volatility of the trade: a trader might be forced out of the positions because of a large, adverse intraday change in the net basis.

[13] At $100,000 face amount per contract, 10,000 contracts is equivalent to the $1 billion face amount of bonds. The conversion factor of the CTD into TYM0 is 0.8006, the repo rate is 1.10%, and there are 116 days from March 6, 2020, to June 30, 2020. Therefore, the conversion-factor weighted and tailed number of contracts is $10{,}000 \times 0.8006/(1 + 1.10\% \times 116/360) = 7{,}978$.

Another way a trader might be forced out of the basis trade is the need to come up with cash. As of March 6, with maintenance margin at $1,275 per contract, the short of 7,978 contracts requires posting a bit more than $10 million to the exchange.[14] With the trade losing money on March 9, the hedge fund needs to make daily settlement payments on the futures contracts and variation margin payments on the term repo positions, which essentially sum to the P&L for the day, that is, to about $3.2 million. In addition, because of increased market volatility, the exchange increased the margin requirement from $1,275 to $1,375 per contract, generating a margin call of nearly $800,000. In total, then the hedge fund has to come up with an additional $4 million on March 9. This is a very large number relative, once again, to the target profit of $2.5 million. These cash requirements also clarify the concreteness of allocated capital, because that capital really does need to be available to meet the cash requirements of the trade. In any case, as the trade makes money after March 9, cash requirements fall, but not as quickly as indicated by the P&L; the exchange increased margin requirements once again on March 11.

A final way a trader might be forced out of a basis trade applies only to those that finance the bond purchases with overnight rather than term repo. In calm markets, rolling over overnight repo is nearly equivalent to term repo. More precisely, if the term repo rate reflects future, unchanging overnight rates, and if financing is plentiful, then there is no important difference in borrowing overnight and borrowing term. In more volatile markets, however, borrowing overnight adds two significant sources of risk.

First, the overnight rate can rise or fall relative to what could have been locked in to term. In March 2020, as shown in Figure 11.13, the overnight repo rate, as represented by SOFR, declined precipitously as the Federal Reserve lowered rates to accommodate market conditions. Bearing this risk actually worked in favor of traders who had not locked in term repo rates – with one caveat to be discussed presently – as they could roll overnight repo at lower and lower rates. Another way to express the advantage of not locking in term rates is to compare the gross and net basis of TYM0 in Figure 11.10. As the repo rate falls, the carry of the CTD increases, which means that its gross basis increases relative to its net basis.

Second, in stressed markets, prime brokers might reduce risk by limiting lending to hedge funds in the basis trade, or, with similar effect, by increasing

[14]Treasury repo does typically require margin, given in Table 10.3 at about 2%. However, because the risks of futures and bond positions in a basis trade are largely offsetting, a prime broker might very well – at least in normal times – not require a hedge fund to post much more than exchange-required futures margin. The text assumes this even though, for some clients, prime brokers will require a cushion above exchange margin.

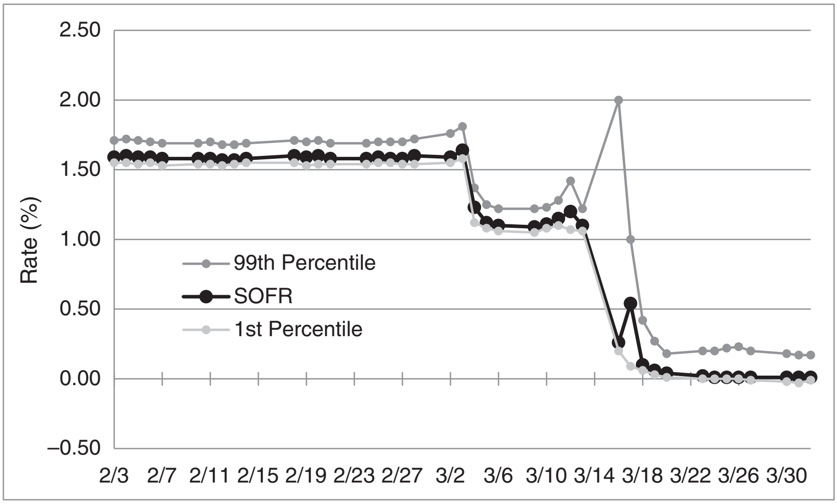

FIGURE 11.13 SOFR and 1st and 99th Percentile Repo Rates, from February 3, 2020, to April 1, 2020.

repo haircuts. In that case, basis traders who did not lock in term repo might have to reduce position size or even exit a position at a loss. Some evidence of this funding risk through March 2020 can be seen from the difference between the 99th percentile repo rate and SOFR in Figure 11.13. Through February, the spread of the highest rate percentile of repo transactions over the median SOFR rate was typically only 10 or 11 basis points. In March, however, that spread was between 20 and 50 basis points on several days and hit a peak of 174 basis points on March 16. These data strongly imply that at least some overnight repo borrowers had trouble borrowing money against Treasuries. Furthermore, basis traders who had not locked in term rates might very well have had to roll over borrowing at rates very much above term rates they might have locked in at an earlier date.

CHAPTER 12

Short-Term Rates and Their Derivatives

Short-term funds trade in *money markets* in significant volumes and at a variety of interest rates. Interest rates vary across currencies, of course, but even within a single currency there are typically several rates that reflect differences in collateral, market participants, and term. This chapter introduces some of the most important short-term interest rates across currencies and describes the global transition away from LIBOR. The chapter then describes a number of interest rate derivatives contracts – one-month SOFR futures, fed funds futures, three-month SOFR futures, Euribor FRAs, and Euribor futures – and explains their use in hedging exposures to short-term interest rates. A detour shows how combining fed funds futures prices with Federal Reserve Board meeting dates is used to imply expectations of the Federal Reserve's target rate and to construct benchmark interest rate curves for general use. The chapter concludes with a brief discussion on the difference between futures and forward rates. Notes on pricing rate forwards and futures on rates with term structure models can be found in the appendix to Chapter 12.

12.1 SHORT-TERM RATES AND THE TRANSITION FROM LIBOR

Since a first trade in 1969,[1] *LIBOR* (*London Interbank Offered Rate*) grew to become the dominant set of reference rates for short-term interbank borrowing. Across a wide variety of loans and derivatives, huge volumes of cash flows were set based on LIBOR. LIBOR rates were published daily for five currencies (CHF, EUR, GBP, JPY, USD) and for seven maturities (overnight, one week, one month, two months, three months, six months, 12 months), although the three-month and to a much lesser extent the one-month terms were the most popular. For each currency, a panel of 16 banks was surveyed

[1]Reuters (2012),"Insight: A Greek Banker, the Shah and the Birth of LIBOR," August 7.

with the question, "At what rate could you borrow funds, were you to do so by asking for and then accepting interbank offers in a reasonable size just prior to 11 a.m. (London time)?" The LIBOR *fixing* for each currency and term was then set equal to a trimmed mean of the survey responses.

In the wake of the quantitative easing policies of central banks after the financial crisis of 2007–2009, the survey methodology began to lose credibility. First, because banks with abundant reserves do not need to borrow reserves from other banks, the volume of interbank transactions fell so dramatically that survey responses became more subjective. Second, from 2012 to 2016, investigations revealed that traders at the panel banks colluded to manipulate LIBOR fixes to their advantage. Some traders were sentenced to prison, and banks collectively paid more than $9 billion in fines.

In 2017, regulators decided to terminate LIBOR and its use by the end of 2021. Furthermore, after consultations with industry representatives, it was decided that financial contracts in the future should reference a set of overnight, "risk-free rate" (RFR) benchmarks, which can be set objectively based on large volumes of transactions data. The phrase "risk-free rates" is somewhat of a misnomer, as only two of the new benchmarks – SARON in CHF and SOFR in USD – are based on collateralized repo rates, while the other three – €STR in EUR, SONIA in GBP, and TONAR in JPY – remain interbank rates. And although all of the new benchmarks are overnight rates, regulators have permitted the continued use of two term rate indexes, after methodological reforms, namely, Euribor in EUR and TIBOR in JPY. Table 12.1 and the following discussions outline the rates used before the transition and the RFRs compliant with post-transition regulatory principles. The table highlights in bold the overnight benchmarks in each currency that are preferred by the regulators.

TABLE 12.1 Global Market Short-Term Interest Rates, Pre- and Post- the Transition Away from LIBOR.

		Post-Transition		
			Interbank (Unsecured)	
Currency	Pre-Transition	Repo	Overnight	Term
CHF	LIBOR	**SARON** (2009)		
EUR	EONIA, Euribor		**€STR** (2019)	Euribor
GBP	LIBOR, SONIA		**SONIA**	
JPY	LIBOR, TIBOR, Euroyen TIBOR		**TONAR** (2016)	TIBOR
USD	fed funds, LIBOR	**SOFR** (2017)	fed funds	Ameribor BSBY

CHF: Swiss franc; EUR: euro; GBP: British pound; JPY: Japanese yen; USD: US dollar.

Swiss Franc (CHF). The CHF benchmark moved from LIBOR to SARON (the *Swiss Average Rate ON*, where *ON* stands for overnight). SARON is a volume-weighted average of the rates on general collateral, fixed income repo transactions. SARON is calculated a few times per day, along with indexes for several other terms (e.g., SAR1W for one week, and SAR3M for three months), but SARON at each day's close is the main benchmark rate.

Euro (EUR). *EONIA (Euro ON Index Average)* was a volume-weighted average of rates on overnight interbank loans, which was computed from transactions supplied by a panel of banks. Over time, the total volume of transactions, along with the number of banks in the panel with significant volumes, had fallen too much for the index to remain tenable. *Euribor (Euro Interbank Offered Rate)* had been a set of five term rates (one week, one month, three months, six months, 12 months) based on submitted quotations by panel banks, like LIBOR, although with a different administrator. As part of the transition away from LIBOR, Euribor was reformed in 2019 to follow a "hybrid methodology," which, to the extent available, uses Level 1 contributions, which consist of wholesale funding transactions of panel banks – 18 at the time of this writing – on a given day and for a given term. To the extent such transactions are not available, Euribor may rely on Level 2 contributions, which are recent transactions at other terms, or, if necessary, on Level 3 contributions, which are relevant transactions from other, closely related markets. Perhaps not surprisingly, given the monetary regime of abundant reserves, a significant fraction of contributions are, indeed, Level 3.[2] *ESTR*, *ESTER*, or *€STR* (*Euro Short-Term Rate*) is the index favored by regulators. It is a volume-weighted trimmed mean of sizable, overnight funding transactions by – at the time of this writing – 32 euro-area banks.

British Pound (GBP). *SONIA (Sterling ON Index Average)* has existed since 1997, has a history of use as a reference rate for derivatives transactions, and was reformed in 2018. Like €STR, the reformed *SONIA* is a volume-weighted trimmed mean of sizable overnight bank funding transactions.

Japanese Yen (JPY). There were two versions of *TIBOR* (*Tokyo Interbank Offered Rate*), both based on submissions by Japanese banks: *Euroyen TIBOR*, representing interbank JPY rates outside Japan (like JPY LIBOR), and *JPY TIBOR* or simply *TIBOR*, representing interbank JPY rates inside Japan. While Euroyen TIBOR has

[2]See, for example, Amor, J. (2021), "The New EURIBOR Gets Through a Challenging 2020," *Funcas SEFO* 10(1), pp. 49-61, January.

been abandoned, TIBOR was reformed in various ways in 2017, including – along the lines of the Euribor reforms – a waterfall of submission *priorities*, from actual interbank *call* or funding transactions, through data from similar or relevant markets, down to "Expert judgment," though this last resort has not yet been necessary. TIBOR is published for five terms: one week, one month, three months, six months, and 12 months. *TONAR* (*Tokyo ON Average Rate*), also called *TONA*, was created in 2016 as a volume-weighted average of transactions in the overnight call market.

US Dollar (USD). fed funds is the overnight rate at which banks trade reserves at the Federal Reserve System. A further description of this market is given below. *SOFR* (*Secured Overnight Financing Rate*) is a trimmed, volume-weighted median of overnight repo rates. Its calculation is described in more detail in Chapter 10. The other USD rates listed in Table 12.1 are discussed presently.

Given the enormous volume of loans and derivatives that had been referencing LIBOR, the transition to new rates, from an operational perspective, was time-consuming and costly. Changes to derivatives markets were organizationally straightforward, in that regulators could push changes through a small number of clearinghouses and large dealers. By contrast, changes in loan markets required action by a large number of smaller banks. Particularly challenging in both markets, however, are *legacy* or existing contracts that reference LIBOR, but are difficult to change for various reasons, ranging from their being in structures that are hard to amend to their somewhat unresponsive counterparties. The resulting backlog of legacy contracts eventually forced regulators to postpone the cessation of a subset of LIBOR indexes. Five USD LIBOR tenors (overnight, one month, three months, six months, and 12 months) will continue to be published until June 30, 2022, for use by certain legacy contracts only. Three GBP and JPY LIBOR tenors (one month, three months, and six months) will continue to be published for legacy contracts for some time in *synthetic* form, which means that, instead of being calculated from survey results, as in the past, these LIBOR rates are now set as some fixed function of the preferred, post-transition benchmarks, namely, SONIA and TONAR, respectively.

More substantive challenges to the transition, particularly in the loan market, arise from two key features of the new benchmarks. First, they are overnight rather than term rates. Second, they are rates with little to no bank credit risk. The challenge of moving from term to overnight rate benchmarks is described in Figure 12.1. Panel a) depicts the historical use of one-month LIBOR as part of a longer-term, floating-rate loan. A borrower of \$1 million from October 1 to November 1, 2021, knows the rate of interest at the start of the period and, therefore, can compute the interest due at the end of the period, in this case, $\$1{,}000{,}000 \times 0.07525\% \times 31/360 = \$64.80$.

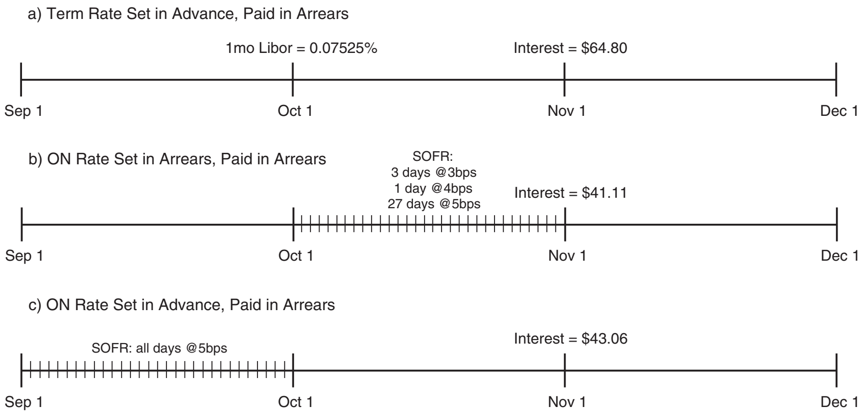

FIGURE 12.1 Examples of Monthly Interest Payment Conventions for a $1 Million Monthly Resetting Floating-Rate Loan, September to November 2021.

In this convention, interest is said to be *set in advance* of the loan period and *paid in arrears* of the loan period. With overnight rather than term benchmarks, however, borrowers typically do not know the final amount of interest at the start of the period. Panel b) shows the most straightforward structure for term loans with an overnight rate benchmark, namely, set in arrears and paid in arrears. Over the 31 days – October 1, 2021, to October 31, 2021 – SOFR was three basis points on three days, four basis points on one day, and five basis points on 27 days. On a $1 million loan, therefore, with daily compounding, the interest payment due at the end of the period is,

$$\$1{,}000{,}000\left[\left(1+\frac{0.03\%}{360}\right)^{3}\left(1+\frac{0.04\%}{360}\right)\left(1+\frac{0.05\%}{360}\right)^{2}-1\right]$$
$$= \$41.11 \qquad (12.1)$$

But the borrower does not know this amount until the end of the period, when all relevant SOFR rates have been observed. Many financial institutions, whose business is the money markets, might be perfectly comfortable with this uncertainty. Nonfinancial corporate borrowers, however, might not be comfortable with uncertain interest payments at the end of one-month, three-month, or longer reset periods.

There exist several operational fixes designed to give borrowers time to compute and make interest payments. A *payment delay* simply gives the borrower an extra day or so to make the payment. A *lockout* allows for earlier computation of interest payments by not using the very last rate settings. For example, a one-day lockout might use the rate for October 30 for both October 30 and October 31. And a *lookback* uses earlier rather than the latest rates. In the present example, the observation period might be shifted from October 1 through October 31 to September 30 through October 30. These minor tweaks, however, do not change the fact that borrowers do not know how much they will pay in interest until sometime near the end of the period.

One proposed solution to this more significant challenge, depicted in Panel c) of Figure 12.1, is to use overnight rates in a set in advance, paid in arrears structure. For example, to determine the payment on November 1, 2021, use the interest rates from September 1 to October 1 (inclusive). Over this period, SOFR was five basis points every day, giving an interest payment of $\$1{,}000{,}000 \times \left[(1 + 0.05\%/360)^{31} - 1\right] = \$43.06$. In this way, the borrower does know the interest due on November 1 as of October 1. The reference interest rates, however, do not correspond to the borrowing period, which may or may not be important depending on the situation and on the structure of any complementary positions.

The most natural solution, from a theoretical perspective, is for the market to develop term rates, based on the benchmark overnight rates, for use in loans to borrowers who want payment certainty. Some such term rates are, in fact, in use for some purposes, including the synthetic versions of GBP and JPY LIBOR mentioned already. And the CME, a major US derivatives clearinghouse, has been encouraged to create term SOFR rates for limited use. For now, however, regulators are discouraging – as a return to the recent weaknesses of LIBOR – the broad use of term rates that are not directly observable from significant volumes of market transactions. In the future, perhaps, the rate derivatives described later in the chapter will trade with enough liquidity to provide acceptable term rates.

The second substantive challenge to the LIBOR transition is that the new RFR benchmarks embed little to no bank credit risk. In the heyday of LIBOR, banks would borrow funds at LIBOR or at a spread to LIBOR and would lend funds to customers at LIBOR plus a spread. Banks were naturally hedged, therefore, against increases in the cost of funds of the banking sector. Some of the new overnight benchmarks are interbank rates, namely, €STR, SONIA, and TONAR, but overnight interbank rates – as opposed to term interbank rates – do not typically fluctuate much with credit conditions until well into a financial crisis. And the other two new benchmarks, SARON and SOFR, are based on repo rates, which, if anything, might fall as credit conditions worsen, as flight-to-quality trades flock to safer lending havens. The question for the transition, therefore, is whether banks will benchmark large volumes of their customer loans to RFRs, and even if they do so in normal times, will they continue to do so in times of stress.

In EUR and JPY, the continued use of term Euribor and TIBOR enables the use of *credit-sensitive rates* (CSRs). In the United States, there are two suites of rates that have been deemed by independent auditors as compliant with international regulatory principles pertaining to benchmarks.

Ameribor (*American Interbank Offered Rate*) is computed for a range of terms from the funding transactions and executable quotes of a large number of small to regional US banks on the American Financial Exchange (AFX). *BSBY* (*Bloomberg Short-Term Bank Yield Index*) is computed for a range of terms from the funding transactions and executable quotes of large banks across the Bloomberg trading platform. Futures contracts are listed on both Ameribor and BSBY rates. *AXI* (*Across-the-Curve Credit Spread Index*) is not listed in Table 12.1, because, at the time of this writing, it had not yet been deemed compliant with regulatory benchmark principles. This index is intended as a credit spread to be added to SOFR, is available for a range of terms, and is based on unsecured bank funding transactions. Ameribor, BSBY, and AXI, from a commercial perspective, face not only the normal challenges in bringing new products to market but also the regulatory headwinds that support broad adoption of SOFR.[3]

12.2 ONE-MONTH SOFR FUTURES

The one-month SOFR futures contract, which trades on the CME (Chicago Mercantile Exchange), is designed to take and hedge exposure to SOFR or to other rates believed to be highly correlated with SOFR. Selected one-month SOFR contracts, along with their prices and rates as of January 14, 2022, are given in Table 12.2. Each ticker is composed of the code "SER," a letter indicating the contract month, and a digit corresponding to the last digit of the contract year. For example, with "G" standing for February, SERG2 is the ticker for the one-month SOFR contract of February 2022. Its current price is 99.935, which corresponds to a percentage rate of 100 – 99.935, or 0.065. Expressing the percentage rate as 100 minus price is just a convention: there is no sense in which the price of a particular bond or other instrument at a rate of 0.065% equals 99.935.[4]

[3]For example, the OCC (Office of the Comptroller of the Currency) noted that "While banks may use any replacement rate [for LIBOR] they determine to be appropriate for their funding model and customer needs, OCC supervisory efforts will initially focus on non-SOFR rates." OCC (2021), "LIBOR Transition: Updated Self-Assessment Tool for Banks," OCC Bulletin 2021-46, October 18. And an executive vice president and general counsel at the Federal Reserve Bank of New York said that "If your firm is moving to a rate other than SOFR, that means extra work for you to make sure you're demonstrably making a responsible decision. Because you are going to be asked about that decision." Held, M. (2021),"Prepare for Landing," Remarks at the ISDA Benchmark Strategies Forum, Federal Reserve Bank of New York, September 15.

[4]The minimum price increment and, therefore, the precision or number of decimal places in the price of each contract, depends on the day of the week of the first day in the settlement month and on the remaining time to contract expiration.

TABLE 12.2 Selected One-Month SOFR Futures Contracts, as of January 14, 2022. Rates Are in Percent.

Ticker	Month	Price	Rate
SERF2	Jan	99.9475	0.0525
SERG2	Feb	99.935	0.065
SERH2	Mar	99.815	0.185
SERJ2	Apr	99.685	0.315
SERK2	May	99.585	0.415

Each contract trades until the last business day of the month in which it expires, known as its *delivery* month. The word "delivery" here is a carry-over from other futures contracts, which, at expiration, require the delivery of an underlying commodity or financial instrument in exchange for some payment. SOFR contracts, however, are *cash settled*; that is, they require no physical delivery at expiration.

The final settlement rate of a one-month SOFR contract equals the average SOFR rate over that month. Each day of the month is included in the average, and non-business days take the rate of the previous business day. For example, the SOFR rate for Friday, February 4, 2022, is counted three times in the February average: once for Friday, February 4; once for Saturday, February 5th; and once for Sunday, February 6th. The final settlement price of a contract equals 100 minus the percentage rate. If, for example, the average SOFR rate over the month of February 2022 is 0.09%, then the final settlement price is 100 – 0.09, or 99.91.[5]

The one-month SOFR contract is scaled to hedge a $5 million 30-day investment. More specifically, the P&L (profit and loss) on one contract from a one-basis-point decrease in the contract's rate is set equal to,

$$\$5{,}000{,}000 \times 0.01\% \times \frac{30}{360} = \$41.67 \tag{12.2}$$

To elaborate on the P&L from buying, holding, and selling this futures contract, consider a trader that buys one February contract on January 14, 2022, at a price of 99.935, which corresponds to a rate of 0.065%. Say further that the price falls to 99.910, or equivalently, that the rate rises to 0.09%, and that the trader either sells the contract or that the contract expires at those levels. Because the trader bought the contract and the price

[5]The Federal Reserve Bank of New York releases the SOFR of a given day on the morning of the next business day, which determines the earliest availability of the monthly average. February 28, 2022, for example, is a Monday, which means that the February average is available on March 1.

fell, the trader loses money. The actual loss is determined with reference to Equation (12.2). Because the contract rate rose from 0.065% to 0.09%, that is, by 2.5 basis points, the trader loses 2.5 times $41.67, or $104.175. If, the price rises instead, the buyer of a contract makes money. More specifically, if the price rises from 99.935 to 99.995, or, equivalently, the rate falls by six basis points, then the P&L from the purchase of one contract is six times $41.67, or $250.02.

Like other futures contracts, one-month SOFR contracts are subject to *daily settlement* payments. Futures prices fluctuate throughout the day, according to market forces, and, at the end of each day, the exchange sets a settlement price, which usually equals the price of the last trade. All market participants then settle up their gains and losses over the day. Those who were long one-month SOFR contracts over the day pay $41.67 for every one-basis-point increase in rates over the day or receive $41.67 for every one-basis-point decrease in rates over the day. Correspondingly, those who are were short over the day receive $41.67 for every one-basis-point increase in rates or pay $41.67 for every one-basis-point decrease in rates.[6]

Because of daily settlement payments, the P&L of a futures position is realized over its holding period rather than all at unwind or at final settlement. In one of the previous examples the February contract is bought on January 14 at a rate of 0.065% and expires or is sold at a rate of 0.005%, for a six-basis-point or $250.02 gain. That gain is realized over time: daily settlement payments are made on days when rates rise and are received on days when rates fall. The sum of all those daily settlement payments, however, corresponds to the overall six-basis-point decline in rates, that is, to an overall profit of $250.02. The exact pattern of gains and losses over the holding period does not matter much when interest rates are very low, as they have been for several years. In higher rate environments, however, the pattern does matter. For winning positions, early gains, which can be reinvested, can be worth significantly more than later gains. Similarly, for losing positions, early losses, which have to be financed, can cost significantly more than later losses. The implications of daily settlement are discussed further presently.

The stage is now set for a simple example of hedging with one-month SOFR contracts. Consider a money market fund that plans to invest $50 million from January 14 to March 14, 2022, in overnight repo. The fund chooses overnight repo because of its liquidity, that is, its ready availability in the contingency of larger than anticipated requests to redeem shares. At the same time, however, the fund decides to hedge the interest rate risk of falling repo rates with one-month SOFR futures. More specifically, to offset losses

[6]For more on the operations and risk management of the exchange and its associated clearinghouse, see Chapter 13.

from falling repo rates, the fund buys one-month SOFR futures contracts, which profit from falling repo rates. The question for the fund then, is which contracts and how many to buy.

Focusing first on hedging the risk of falling repo rates over the month of February, the fund can buy February contracts. Because each contract is scaled to hedge a \$5 million 30-day investment, while the fund is investing \$50 million over the 28-day month of February, the fund needs to buy,

$$\frac{\$50{,}000{,}000 \times 28}{\$5{,}000{,}000 \times 30} = 10 \times \frac{28}{30} = 9.33 \qquad (12.3)$$

contracts. Under a few simplifying assumptions, discussed presently, the performance of this hedge is described in Table 12.3. The first column gives the average realized rate over the month of February. The second column gives the investment proceeds from earning simple interest in repo over the month of February. At an average repo rate of 0.09%, for example, the fund earns \$50 million $\times$ 0.09% $\times$ 28/360 = \$3,500. The third column gives the P&L from being long 9.33 contracts. Taking again the realized average rate of 0.09%, which is 2.5 basis points above the purchase rate of 0.065%, the long futures position loses $9.33 \times 2.5 \times 41.67 = \$971.95$. The fourth column gives the net proceeds, that is, the sum of the investment proceeds and the P&L from the SOFR futures hedge. The hedge works well in the sense that, no matter what the realized average rate, net proceeds are about the same. To elaborate, buying SOFR futures at a rate of 0.065% approximately locks in a net P&L of \$2,528. If the realized repo rate is less than 0.065%, investment proceeds are less, but futures P&L profits compensate for the difference. If the realized repo rate is greater than 0.065%, investment proceeds are greater, but futures P&L losses offset the difference. The reader can verify that the hedge works perfectly, that is, net proceeds always equal \$2,527.78, under the contrafactual assumptions that: i) the fund buys exactly 9 1/3 contracts, which is the exact solution of Equation (12.3); and ii) the P&L of the futures contract is exactly \$41 2/3 per basis point, which is the exact expression in Equation (12.2).

TABLE 12.3 Hedging a \$50 Million Overnight Repo Investment with 9.33 One-Month February SOFR Futures Contracts, Purchased at a Rate of 0.065%, as of January 14, 2022. Rates and in Percent. Other Entries Are in Dollars.

Realized Avg. Rate	Investment Proceeds	Futures P&L	Net Proceeds
0.050	1,944.44	583.17	2,527.62
0.065	2,527.78	0.00	2,527.78
0.090	3,500.00	−971.95	2,528.05

The depiction of the hedge in Table 12.3 is simplified in several ways. First, the fund has to buy a whole number of contracts, that is, nine (or 10) rather than 9.33. This is less of an issue, of course, for larger hedges, and another practical solution to this rounding problem is discussed presently. Second, the table ignores interest on daily settlement payments, which is negligible when rates are very low and the time horizon is short. Third, it is assumed that investment proceeds earn simple interest, whereas interest is more likely to compound. This effect, too, is small when rates are low and the horizon short. Fourth, the table assumes that the average repo rate that determines the fund's investment proceeds is exactly the same as the average SOFR rate that determines the final settlement rate of the futures contract. While extremely reasonable, this assumption may not hold exactly in practice. SOFR on any day is the trimmed median rate of all transactions that day. The fund, on the other hand, executes its own repo transactions each day at a particular time with a particular dealer.

Stepping back to the full example, the money market fund wants to hedge its overnight repo investments from January 14 to March 14. The hedge over February is described above. Turning to the hedge over the month of January, note that SOFR observations for the first 13 days of January have already been set. Therefore, a one-basis-point change to repo rates over each of the remaining 18 days of January moves average SOFR for the month by only 18/31 basis points. Similarly, the fund is investing in repo over only 18 days in January. The correct number of contracts for the hedge is, therefore,

$$\frac{\$50{,}000{,}000 \times 0.01\% \times \frac{18}{360}}{\$5{,}000{,}000\left(0.01\% \times \frac{18}{31}\right) \times \frac{30}{360}} = 10 \times \frac{31}{30} = 10.33 \tag{12.4}$$

In words, because the number of remaining days in the month affects the investment proceeds and the futures contract in the same way, the fact that the hedge starts in the middle of the month can be ignored. Therefore, just as in Equation (12.3) for the February hedge, the number of contracts for the January hedge is the investment of \$50 million for 31 days divided by the contract's effective hedge of \$5 million for 30 days, which equals 10 times 31/30, or 10.33. Table 12.4 illustrates the January hedge, noting that SOFR over each of the first 13 days of January was five basis points. The format of the table is similar to that of Table 12.3, but the first column is the average repo rate over the whole month, which determines the futures P&L, while the second column is the average repo rate over the last 18 days of the month, which determines the investment proceeds. From the first row, for example, the first 13 days at five basis points and the last 18 days at 2.85 basis points give an average over the month of $(5 \times 13 + 2.85 \times 18)/31 = 3.75$ basis points.

TABLE 12.4 Hedging a $50 Million Overnight Repo Investment with 10.33 One-Month January SOFR Futures Contracts, Purchased at a Rate of 0.0525%, as of January 14, 2022. SOFR Was 0.05% on Every Day from January 1 to 13. Rates Are in Percent. Other Entries Are in Dollars.

Jan 1-31 Realized Avg. Repo	Jan 14-31 Realized Avg. Repo	Investment Proceeds	Futures P&L	Net Proceeds
0.0375	0.0285	711.81	645.68	1,357.48
0.0525	0.0543	1,357.64	0.00	1,357.64
0.0775	0.0974	2,434.03	−1,076.13	1,357.90

Table 12.4 shows that the January hedge performs as expected. The January SOFR futures price is at a rate of 0.0525%, which implies a rate of 0.0543% over the last 18 days of the month. The money market fund can lock in the investment proceeds at this rate, that is, $50 million × 0.0543% × 18/360, or about $1,358. At lower realized rates, investment proceeds are lower, but futures gains compensate. At higher realized rates, investment proceeds are higher, but are given back through futures losses. Note that, as in the previous table, the hedge would work exactly for 10 1/3 contracts and a futures profit or loss of $41 2/3 per basis point.

The discussion of this January hedge reveals, more generally, that hedges with one-month SOFR futures from any date within a month to the end of that month require no special handling. Hedge ratios are computed as if the entire month were being hedged, and, consequently, hedge ratios do not change over the course of the contract month. Applying this conclusion to the February part of the hedge, described earlier, the fund maintains its initial long of 9.33 contracts throughout the entire month of February, after which time those contracts expire.

The remaining piece of the overall money market fund hedge is using the March contract to hedge repo rates from March 1 to 13, which rates determine the fund's investment proceeds from March 1 to the assumed horizon date of March 14. Proceeding in the same way as for January and February, the fund would hedge its $50 million investment over 13 days with contracts that hedge $5 million over 30 days by buying 10 times 13/30, or 4.33 March contracts. With the rate of the March contract at 0.185% as of January 14, this approach would seemingly lock in a repo investment rate of 0.185%. It is left to the reader to construct a table like Tables 12.3 and 12.4 to illustrate this point. The March hedge is actually more complicated than the January and February hedges, however, for two reasons. First, 4.33 contracts are appropriate from January 14 until before the setting of SOFR on March 1, but not for the rest of March. Consider, for example, the hedge on March 4, before the setting of SOFR on that day. There are 10 days left of exposure to the investment – March 4 to March 13, inclusive of both dates – and 28 days left of exposure to the contract – March 4 to March 31, inclusive of both dates. The hedge ratio, therefore, is given by,

$$\frac{\$50{,}000{,}000 \times 0.01\% \times \frac{10}{360}}{\$5{,}000{,}000 \left(0.01\% \times \frac{28}{31}\right)} = 10 \times \frac{31}{28}\frac{10}{30} = 3.69 \tag{12.5}$$

In short, the appropriate number of contracts falls gradually over March from 4.33 before the setting of repo on March 1 to zero on March 14. More generally, the number of one-month SOFR contracts hedging exposure from the beginning of the month to sometime within the month declines over the course of the month.

The second and more serious problem with the proposed March hedge is its implicit assumption that repo rates in the first 13 days of March move in parallel with rates over the whole month. In fact, the Federal Reserve announces on March 16, 2022, whether it is or is not raising its short-term policy rates. The problem for the money market fund in the example, then, is the scenario in which the market raises its expectation for a policy rate increase on March 16. In that case, rate expectations for the second half of March rise; the SOFR March contract rate rises and its price falls; the money market fund loses money on its futures hedge; but, because repo rates do not rise before the policy change, in the first half of March, fund investment proceeds do not rise.

In this particular example, there is a reasonable solution to the problem of the March hedge: hedge the exposure to repo interest rates from March 1 to March 13 with February SOFR contracts! The prior scheduled announcement of Federal Reserve policy rates is on January 26. Much of the fund's exposure to repo rates in the first part of March, therefore, is captured by February SOFR contracts, which incorporate expectations and realized policy decisions from the January meeting. Therefore, instead of purchasing 4.33 March contracts at the start of the hedge, on January 14, the fund can purchase an additional 4.33 February contracts. Of course, at the end of February, when the contracts expire, the fund is faced with another decision. It can essentially roll the hedge from February to March contracts, by purchasing 4.33 March contracts as the 4.33 February contracts expire. Or the fund might decide that the basis risk between changes in rates over the first part of March relative to changes in rates over the whole of March is large relative to the risk of repo rates moving much before the scheduled March policy meeting. In that case, the fund might choose not to hedge. In any case, the strategy of *stacking* the fund's March risk into February contracts is not a cure-all: some economic and financial surprises might affect repo rates in February, but not in the first part of March, while other surprises might affect repo rates in the first part of March, but not in February. All in all,

TABLE 12.5 Two Hedging Strategies for a \$50 Million Overnight Repo Investment with One-Month SOFR Futures Contracts, as of January 14, 2022. There Is a Scheduled Federal Reserve Target Rate Policy Announcement on March 16, 2022.

Ticker	Contract Month	Month-by-Month Number of Contracts	Stacking Mar into Feb Number of Contracts
SERF2	Jan	$10.33 = 10 \times \frac{31}{30}$	10.33
SERG2	Feb	$9.33 = 10 \times \frac{28}{30}$	13.67
SERH2	Mar	$4.33 = 10 \times \frac{13}{30}$	0.00
Total		24.00	24.00

however, the risk of rates changing over the policy meeting is likely the most important consideration, which argues for stacking into February contracts.

By way of summary, Table 12.5 describes the overall hedge strategies of the money market fund in this example. The \$50 million amount being hedged and the number of days in each month, set against the \$5 million 30-day SOFR futures contract, give the number of contracts in the "Month-by-Month" strategy. Avoiding the basis risk of hedging the first part of March with all of March, and relying on the assumption that almost all repo rate variability arises from changes to policy target rates, the "Stacking Mar into Feb" strategy uses February contracts to hedge the risk from the first part of March.

The notion of stacking contracts with risks from other periods allows the discussion to circle back to the question of dealing with a fractional number of contracts. Table 12.5 shows that the total number of contracts required in either hedge is 24. To avoid fractional contracts, then, while keeping overall exposure correct, stack the fractional amounts from some months into others. In the Month-by-Month strategy, for example, the fund might divide the total 24 contracts into 10 January contracts, 10 February contracts, and 4 March contracts. In the "Stacking Mar into Feb" strategy, the fund might – to avoid too much idiosyncratic February risk – buy 11 January contracts and 13 February contracts. Any stacking strategy does expose the fund, however, to any idiosyncratic changes in rates that appear in one month, but not in another.

12.3 FED FUND FUTURES

Banks that have accounts at the Federal Reserve system can trade these *fed funds* with each other; that is, they can borrow from and lend to each other, at market-determined rates, in the *fed funds market*. These trades are predominantly for overnight borrowing and lending, and the *effective fed funds* rate, calculated each day, is the volume-weighted average of the overnight rates on fed funds traded that day.

Before the financial crisis of 2007–2009, under the monetary policy regime of scarce reserves, domestic commercial banks actively traded in the fed funds market to manage their individual shortages or surpluses of reserves. Furthermore, the Federal Reserve conducted monetary policy by adding or subtracting reserves from the system so as to keep effective fed funds at or very close to a policy *target rate*. More recently, however, in the regime of abundant reserves, domestic commercial banks do not need to borrow funds from each other; they face various regulatory incentives to hold reserve balances; and they earn interest on reserves held at the Federal Reserve. The resulting transformation of the fed funds market has been dramatic, both in terms of a significant decline in volumes and in terms of a change in the composition of market participants. Lending is now dominated by Federal Home Loan Banks (FHLBs) – that do not earn interest on their accounts at the Federal Reserve – and borrowing is now dominated by US subsidiaries of foreign banks – that typically do not take US deposits and are not constrained by the full range of US bank regulations. The fed funds market has thus become a means through which FHLBs can earn interest on their surplus cash in the Federal Reserve system, and foreign banks can earn a spread by borrowing from the FHLBs and depositing those borrowed funds into interest-bearing accounts at the Federal Reserve.[7]

While the Federal Reserve no longer changes the quantity of reserves in the system to influence short-term rates, it uses other policy tools to keep effective fed funds within a target range, defined by an upper and lower limit. Figure 12.2 graphs these upper and lower limits, along with the effective fed funds rate itself, from January 2010 to January 2022. Effective fed funds is somewhat volatile, but nearly always within the policy target range.

The fed funds effective rate is of interest to participants in the fed funds market; to those who borrow and lend at a spread over fed funds effective; to those who trade derivatives tied to fed funds effective; and to anyone taking positions on future Federal Reserve policy. For these groups, fed fund futures can be used to manage exposures to fed funds effective. These futures contracts can be explained concisely here, because one-month SOFR futures, described in the previous section, are designed along the same lines. The final settlement rate of a fed fund futures contract in a particular delivery month is the average of fed funds effective over that month, with non-business days

[7]For additional institutional details, see, for example, Craig, B., and Millington, S. (2017), "The Federal Funds Market Since the Financial Crisis," *Economic Commentary* 2017–07, Federal Reserve Bank of Cleveland, April 5.

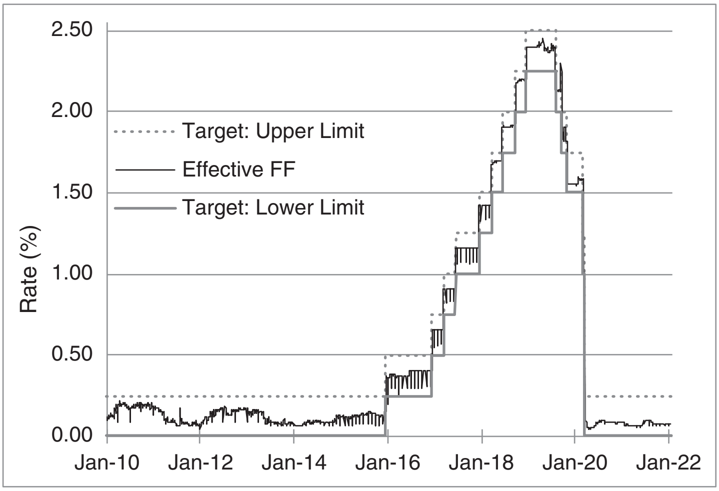

FIGURE 12.2 Fed Funds Effective Rate and Target Range, January 2010 to January 2022.

taking on the previous business day's rate. Like one-month SOFR futures contracts, price is quoted as 100 minus the percentage rate; contracts are subject to daily settlement; and contracts are scaled such that a one-basis-point change in the rate results in a daily settlement flow of \$41.67. With all of these similarities to one-month SOFR contracts, hedging with fed fund futures proceeds along the same lines as in the previous section. The text continues, therefore, with the use of fed fund futures to infer market expectations of Federal Reserve rate policy decisions.

Table 12.6 lists fed fund futures contracts, prices, and rates as of two dates: October 1, 2021, and January 17, 2022. The rightmost column of the table lists the meeting dates of the Federal Open Market Committee (FOMC), which, on the second day of each set of meetings, announces any monetary policy changes, including changes to the federal funds target rate. As explained in Chapter 8, forward interest rates are a combination of expected rates in the future and a risk premium, but the risk premium over a year is likely to be relatively small. Also, as explained later in this chapter, near-term forward rates and futures rates are very similar. Therefore, roughly speaking, the one-month fed fund futures rates in Table 12.6 can be viewed as market expectations of future monthly average effective fed funds. In this light, the main lesson from the table is that expectations of Federal Reserve policy changed significantly from October 1, 2021, to January 17, 2022. As of October 1, fed fund futures anticipated effective fed funds rising from eight basis points to 28 basis points by the end of 2022. As of January 17, 2022, however, by which time surging inflation

TABLE 12.6 Prices and Rates of Selected Fed Funds Futures Contracts, as of October 1, 2021, and January 17, 2022, and 2022 FOMC Meeting Dates. Rates Are in Percent.

		10/1/2021		1/17/2022		FOMC
Ticker	**Month**	**Price**	**Rate**	**Price**	**Rate**	**Meeting**
FFV1	Oct (2021)	99.920	0.080			
FFX1	Nov	99.920	0.080			11/2-3
FFZ1	Dec	99.920	0.080			12/14-15
FFF2	Jan (2022)	99.925	0.075	99.9175	0.0825	1/25-26
FFG2	Feb	99.920	0.080	99.9050	0.0950	
FFH2	Mar	99.920	0.080	99.7800	0.2200	3/15-16
FFJ2	Apr	99.920	0.080	99.6450	0.3550	
FFK2	May	99.915	0.085	99.5500	0.4500	5/3-4
FFM2	Jun	99.900	0.100	99.4500	0.5500	6/14-15
FFN2	Jul	99.885	0.115	99.3450	0.6550	7/26-27
FFQ2	Aug	99.865	0.135	99.2800	0.7200	
FFU2	Sep	99.845	0.155	99.2350	0.7650	9/20-21
FFV2	Oct	99.800	0.200	99.1450	0.8550	
FFX2	Not	99.775	0.225	99.0650	0.9350	11/1-2
FFZ2	Dec	99.720	0.280	99.0050	0.9950	12/13-14

had begun to worry Federal Reserve officials, fed fund futures anticipated effective fed funds rising from 8.25 basis points to 99.5 basis points by the end of 2022.

Many market participants, in fact, use fed fund futures rates together with the FOMC meeting dates to build a market term structure of fed funds effective. The key assumption is that the Federal Reserve changes its target rate only at its scheduled meetings. This has been a good description of recent history, with only rare exceptions, like after the terrorist attacks on September 11, 2001, during the financial crisis in 2008, and in response to the COVID pandemic and economic shutdowns in March 2020. In any case, given this assumption, the idea is to find a set of forward rates, from one meeting to the next, that are more or less consistent with fed fund futures prices.

To illustrate the procedure, focus on January 17, 2022. Note that the fed funds effective rate was eight basis points from January 1 to January 17 (not shown in the table). The FOMC could announce a new target rate the afternoon of January 26 and, because there is no FOMC meeting in February, the rate set by the FOMC on January 26 will be the average fed funds effective rate in February. But the February contract rate is 0.095%. Hence, the market's expectation of fed funds effective after the January 26 meeting is 0.095%.[8] To flesh out the meaning of this expectation, assume further, consistent with recent history, that the Federal Reserve will change the target rate by either zero or 25 basis points. Then, given a post-meeting expected rate of 0.095%, the implied probability, p, of a 25-basis-point increase in rate on January 26, from 8 to 8+25 or 33 basis points, is given by,

$$\begin{aligned} 0.095\% &= p \times (0.08\% + 0.25\%) + (1 - p) \times 0.08\% \\ p &= 6\% \end{aligned} \tag{12.6}$$

In words, with a 6% probability of the Federal Reserve increasing fed funds effective from eight to 33 basis points, and a 94% probability of leaving fed funds effective at eight basis points, the expected fed funds effective rate is 9.5 basis points, as implied by the February contract price.

Moving to the next meeting, because there is a possible rate target announcement on March 16, but no meeting in April, the April futures rate of 0.355% indicates a market expectation of 0.355% after the announcement on March 16. Implying the results of the May meeting is more complex, because there is a meeting in June. Hence, to calculate the market's expectation of the May meeting, let f^{May} be expected fed funds effective after the May meeting, which takes effect starting and including May 5. The May fed futures rate of 0.450%, therefore, is an average of four days at the expectation of 0.355% coming out of the previous meeting, and 27 days at f^{May}. Mathematically,

$$\begin{aligned} 0.450\% &= \frac{4 \times 0.355\% + 27 \times f^{May}}{31} \\ f^{May} &= 0.4641\% \end{aligned} \tag{12.7}$$

Continuing along these lines, the expected fed funds effective rate can be computed for each date. Figure 12.3 shows the results of this exercise, as of October 1 and as of January 17, using fed fund futures rates from Table 12.6. This figure graphically conveys the same message as the futures rates themselves, namely, that the market revised its rate expectation upward from October 1 to January 17. The figure, however, relying on the meeting dates, explicitly shows the path of expected rates.

This section concludes with an aside about curve building, that is, about mathematically representing the term structure of interest rates. Putting aside

[8]There is another way to back out rate expectations of the January meeting, namely, to match the January futures rate as an average of 26 days at eight basis points and five days at the expected result of the January meeting. The purpose of the text, however, is merely to illustrate the procedure, not to recommend a specific approach for this particular point in time.

the small differences between short-term futures and forward rates, discussed next, Figure 12.3 can be read as a *flat-forward-rate* representation of the term structure. In other words, the term structure of forward rates is represented as a sequence of constant forward rates. While this structure may seem odd from a purely economic perspective, because forward rates might be expected to change smoothly over time, flat forwards often represent all available market information. Fed fund futures, for example, provide only one rate per month. Furthermore, while flat forwards are discontinuous; that is, they jump at particular terms, they lead to relatively smooth spot rates and discount factors. To illustrate, using the relationships from Chapters 1 and 2, Figure 12.4 constructs weekly spot rates, discount factors, and forward rates from the flat forwards of Figure 12.3. These curves can reasonably be used to discount cash flows from other short-term instruments of equivalent credit risk, or as a base curve above which spreads can be added to discount riskier cash flows.

In building curves, there is always a trade-off between stability and smoothness. Flat forwards are very stable. Small changes in fed fund futures rates, for example, move flat forwards and corresponding spot rates or discount factors by small amounts and in sensible ways. Smooth representations of interest rate curves, however, like cubic splines, which are not described in this book, are notoriously unstable. Small changes in market rates in one part of the curve often lead to larger changes in other, relatively distant parts of the curve. Current industry practice favors more stable methodologies, like flat forwards, combined with an effort to collect as many market data points as possible, particularly in the short end of the curve.

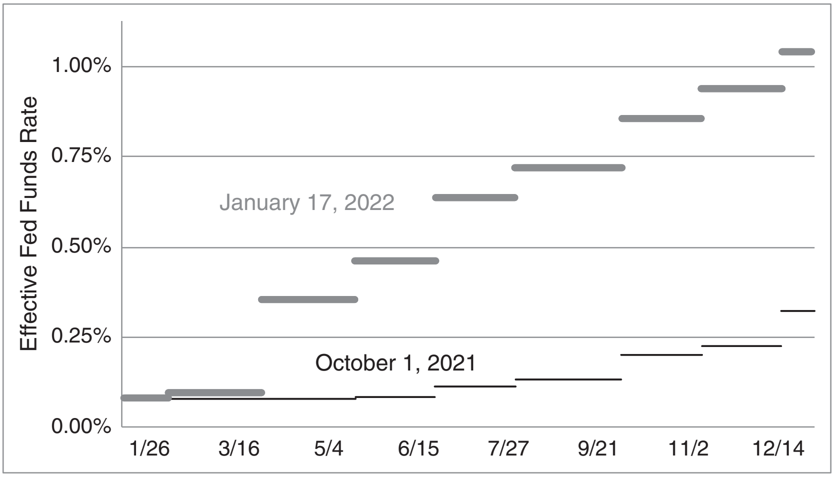

FIGURE 12.3 Implied 2022 fed funds Effective Rates from Fed Fund Futures, as of October 1, 2021, and January 17, 2022.

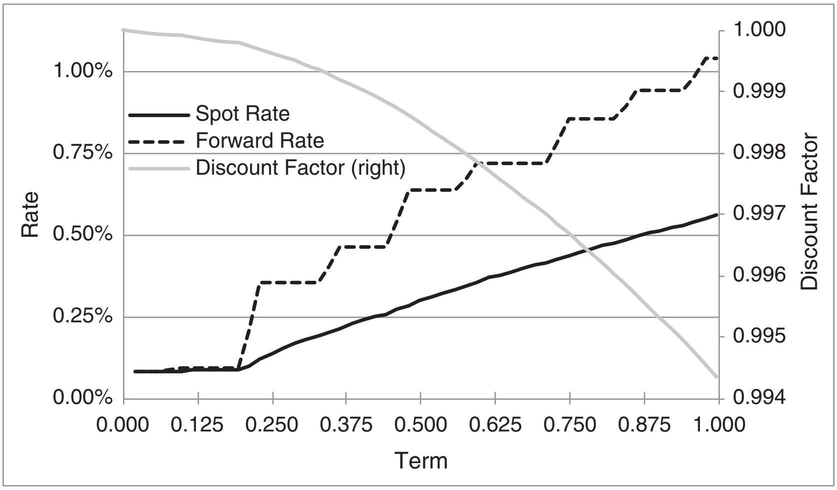

FIGURE 12.4 A Term Structure of Weekly Interest Rates from Fed Fund Futures, as of January 17, 2022.

In general, the more market data, the less significant the discontinuities or jumps in the flat forward curve. Jumps between meeting dates, however, as in Figure 12.3, are actually a feature of the methodology, as they capture the financial reality of short-term rates jumping after FOMC meetings.

12.4 THREE-MONTH SOFR FUTURES

Like one-month SOFR futures, three-month contracts are designed to hedge exposure to SOFR and are subject to daily settlement. There are three differences, however. First, and most obviously, three-month contracts hedge exposure over three months, with the exact determination of dates described presently. Second, the settlement rate of the three-month contract is based not on average SOFR over the reference period, but rather on daily compounded SOFR. Third, the three-month contract is scaled to hedge a \$1 million 90-day investment, which means that the P&L on one contract from a one-basis-point decrease in the final settlement rate is set equal to,

$$\$1{,}000{,}000 \times 0.01\% \times \frac{90}{360} = \$25 \tag{12.8}$$

Table 12.7 lists the first four three-month contracts, along with their prices and rates, as of January 14, 2022. The first three letters of the ticker, “SFR,” denote the three-month SOFR contract, while the following letter and digit, as for other futures contracts, denote the delivery month

TABLE 12.7 Selected Three-Month SOFR Futures Contracts, as of January 14, 2022. Rates Are in Percent.

Ticker	Delivery Month	Start Date (inclusive)	Last Trade Date	Valuation Date	Price	Rate
SFRZ1	Dec	12/15/2021	03/15/2022	03/16/2022	99.940	0.060
SFRH2	Mar	03/16/2022	06/14/2022	06/15/2022	99.640	0.360
SFRM2	Jun	06/15/2022	09/20/2022	09/21/2022	99.350	0.650
SFRU2	Sep	09/21/2022	12/20/2022	12/21/2022	99.125	0.875

and last digit of the year. The "Start Date" of the reference period is the *IMM* (International Monetary Market) date corresponding to the contract month, where IMM dates are defined as the third Wednesday of March, June, September, and December. The "Last Trade Date" of a contract, which is also the last day of the reference period, is the business day before the next IMM date. And the "Valuation Date," which is the day of the last daily settlement payment, is the business day after the last trade date. For the June three-month SOFR contract, for example, the reference period for the calculation of the final settlement rate is from the June IMM date, or June 15, 2022, inclusive, to the September IMM date, or September 20, 2022, inclusive. The contract stops trading on September 20, 2022, and the last daily settlement payment is on September 21, 2022, which is the date that SOFR for September 20, 2022, is published.

Let r_i denote SOFR for date i, and let D denote the total number of days in the reference period. As in the cases of one-month SOFR futures and fed fund futures, SOFR for any non-business day is taken as the value of SOFR on the previous business day. The final settlement rate of a three-month SOFR futures contract, R, is then defined such that,

$$1 + \frac{RD}{360} = \left(1 + \frac{r_1}{360}\right)\left(1 + \frac{r_2}{360}\right)\cdots\left(1 + \frac{r_D}{360}\right) \tag{12.9}$$

$$R = \left[\left(1 + \frac{r_1}{360}\right)\left(1 + \frac{r_2}{360}\right)\cdots\left(1 + \frac{r_D}{360}\right) - 1\right]\frac{360}{D} \tag{12.10}$$

The right-hand side of Equation (12.9) gives the value of an investment of one unit of currency, compounded daily at realized SOFR rates, from days one to D. The left-hand side is the value of an investment of one unit of currency at the term rate R for D days. Equating the two sides of this equation, therefore, means that R is the realized term rate that summarizes daily compounded SOFR over the reference period. Solving for R gives Equation (12.10). To illustrate with an extremely simple example, note that there are 98 days in the reference period of the June three-month SOFR contract. Assume that SOFR over the 42 days from June 15 to July 26, inclusive of both, is 0.51%, while SOFR over the 56 days from July 27 to September 20, inclusive of both, is 0.76%. (Note that these two periods correspond to periods between FOMC meetings listed in Table 12.6.) Then, using Equation (12.10),

$$R = \left[\left(1 + \frac{0.51\%}{360}\right)^{42}\left(1 + \frac{0.76\%}{360}\right)^{56} - 1\right]\frac{360}{98}$$

$$R = 0.653\% \tag{12.11}$$

The final settlement price would be 100 minus the percentage rate, or 100 – 0.653 = 99.347. This price and the rate in Equation (12.11) are, as it turns out, nearly equal to the price and rate of the June contract given in Table 12.7.

To illustrate hedging with three-month SOFR futures, consider the following simple example. As of January 14, 2022, a company plans to borrow $10 million from its bank for the 98 days between June 15, 2022, to September 21, 2022, at daily compounded SOFR plus a spread. To hedge the risk that SOFR is higher over that time period, the company can sell June contracts. Because the loan is for $10 million over 98 days, while each contract is scaled to a $1 million 90-day loan, the number of contracts to be sold is,

$$\frac{\$10{,}000{,}000 \times 98}{\$1{,}000{,}000 \times 90} = 10.89 \tag{12.12}$$

The top panel of Table 12.8 describes the performance of the hedge in three interest rate scenarios. Consider the scenario in which the realized contract rate turns out to be 0.35%. By the definition of that rate in Equations (12.9) and (12.10), the interest in this scenario on the SOFR component of the bank loan (i.e., ignoring the spread) over the 98-period ending September 21, is $-\$10{,}000{,}000 \times 0.35\% \times 98/360 = -\$9{,}527.78$. The P&L from the hedge, that is, from selling 10.89 contracts at 0.65% that expire at 0.35%, for a loss of 30 basis points per contract, is $-30 \times \$25 \times 10.89 = -\$8{,}167.50$. Ignoring daily settlement payments, this panel assumes that this futures P&L is realized on the June contract's valuation date, which is also September 21. Summarizing this scenario, then, the company's relatively low interest cost and its loss on its futures hedge gives a total obligation of about $17,695. In the scenario in which the final settlement rate is 0.65%, the interest cost is about $17,694, and the futures P&L is zero. And if the final settlement rate is 0.95%, then the realized interest cost is relatively high, but, offset by futures gains, again giving a net cost of about –$17,694. Hence, with the three-month SOFR futures hedge, the interest cost is successfully locked in at 0.65%, that is, the rate at which the futures contracts are initially sold. (The net amount would be exactly $17,694.44

TABLE 12.8 Hedging the Interest Cost of Borrowing $10 Million from June 15, 2022, to September 21, 2002, as of January 14, 2022, with June Three-Month SOFR Contracts, Priced Initially at 0.650%. The Hedge, Not Tailed, Sells 10.89 Contracts. The Tailed Hedge Sells 10.84 Contracts. Rates Are in Percent. Other Entries in Dollars.

Final Settle-ment Rate	Loan Interest	Futures P&L	Net
Futures P&L Realized at End – Hedge Not Tailed			
0.35	−9,527.78	−8,167.50	−17,695.28
0.65	−17,694.44	0.00	−17,694.44
0.95	−25,861.11	8,167.50	−17,693.61
Futures P&L Realized at Start – Hedge Not Tailed			
0.35	−9,527.78	−8,187.35	−17,715.13
0.65	−17,694.44	0.00	−17,694.44
0.95	−25,861.11	8,221.38	−17,639.73
Futures P&L Realized at Start – Hedge Tailed			
0.35	−9,527.78	−8,149.76	−17,677.54
0.65	−17,694.44	0.00	−17,694.44
0.95	−25,861.11	8,183.64	−17,677.48

in every scenario, by the way, if the number of contracts in Equation (12.12) were taken to more decimal places.)

The second and third panels of the table return to the discussion earlier in the chapter on the impact of daily settlement payments. The second panel illustrates the issue with the extreme assumption that, immediately after the June contracts are sold on June 14 at a rate of 0.65%, the futures rate jumps to its final settlement value and remains at that level until the end of the loan period on September 21. In the 0.35% scenario, the P&L of −$8,167.50 is fully realized with the daily settlement payment at the close of business on June 14 and has to be financed over the 250 days to September 21 at 0.35%. This brings the total futures loss to $-\$8,167.50(1+0.35\%\times 250/360) = -\$8,187.35$. In the 0.65% scenario, there is no futures P&L, so nothing changes from the first panel. And in the 0.95% scenario, the futures gain of $8,167.50 is fully realized on June 14 and can be reinvested at 0.95% to September 21, bringing the total gain to $8,221.38. The sums of each of these accumulated P&L numbers and the respective loan interest numbers in the second panel of the table give three net quantities that are no longer virtually the same. The cost of financing the daily settlement loss in the first scenario and the interest on the daily settlement gain in the third scenario result in net positions that are not so well hedged as in the first panel.

The differences are not particularly large, because rates are low. But the text continues by describing how to adjust the hedge: interest rates may eventually return to higher levels.

The second panel of Table 12.8 reveals that, incorporating daily settlement payments, selling 10.89 contracts is too many. When rates fall and futures P&L is negative, financing costs drive overall losses too high. And when rates rise and futures P&L is positive, interest earned drives overall gains too high. Roughly speaking then, the daily settlement feature grows any P&L from a futures contract today into that P&L × $(1 + R_0 D/360)$ as of the expiration date, where R_0 is the futures rate today. To offset this growth, a correction used in practice, called *tailing the hedge*, is simply to divide the number of contracts by $(1 + R_0 D/360)$. In the present example, tail the hedge by reducing the number of contracts from 10.89 to $10.89/(1 + 0.65\% \times 250/360) = 10.84$. The third panel shows the P&L of this tailed hedge. In the 0.95% scenario, for example, the futures P&L of $30 \times \$25 \times 10.84 = \$8,130.00$, which is assumed to be realized immediately, is invested at a rate of 0.95% to grow to a total gain of $\$8,130.00(1 + 0.95\% \times 250/360) = \$8,183.64$. All in all, comparing the net quantities in the second and third panels of the table, the tail does reduce the variance of the outcome. The correction is not perfect, of course, because the tail reduces the number of contracts by today's futures rate, whereas daily settlement grows by the rate in the relevant horizon. Note also that the tail needs to be adjusted over time, as the prevailing futures rate changes and as the number of days to contract expiration falls.

The hedging example of this section is very simple in that the company's borrowing dates coincided exactly with the reference period of the June futures contract. If, instead, the company were borrowing for the same number of days, but over a different calendar period, say from April 15 to July 22, the futures hedge still calls for about 11 contracts, but split between March and June. With no view on the relative amount of interest rate risk over the 61 days covered by the March contract (April 15 to June 14, inclusive of both) and the 37 days covered by the June contract (June 15, inclusive, to July 22, not inclusive), the most straightforward split is 11 times 61/98, or about seven in March, and 11 times 37/98, or about four in June.

12.5 EURIBOR FORWARD RATE AGREEMENTS AND FUTURES

As mentioned earlier, Euribor is a set of term rates that continue to be used after the LIBOR transition. In particular, the three-month Euribor quote on a particular date represents the rate on a three-month interbank loan that settles in two business days. Furthermore, a term of "three months" in this market refers to the same calendar day three months later. Three-month Euribor quoted on June 13, 2022, therefore, represents the rate earned over the 92 days from June 15, 2022, to September 15, 2022.

Forward rate agreements or *FRAs* are over-the-counter agreements that are used to hedge the interest rate risk of future borrowing or lending. Table 12.9 gives the terms, as of January 14, 2022, of a €10 million notional amount Euribor FRA at a fixed rate or contract rate of −0.47% from June 15 to September 15, 2022. The *borrower* or *fixed-rate payer* agrees to pay the *lender* or *fixed-rate receiver* on June 15 the quantity listed in the table. The fixed rate of a FRA is set equal to the Euribor forward rate at the time of the trade, as defined in Chapter 2. In this example, this means that, as of January 14, market participants are willing to commit to borrow or lend from June 15 to September 15 at a rate of −0.47%. To see that the FRA is fairly priced at this fixed rate, note that a commitment to borrow or lend for three months on June 15 at the then-prevailing market rate, R, is also fair. Therefore, committing to lend on June 15 at −0.47% and to borrow on June 15 at R is fair, and the resulting September 15 interest payments from those commitments sum to €1 million × (92/360) × $(-0.47\% - R)$. Finally, the present value of that sum as of June 15 equals the quantity in Table 12.9. Hence, the FRA described is fairly priced.

To illustrate how FRAs can be used for hedging, consider a corporation that, as of January 14, 2022, has a line with its bank to borrow €10 million from June 15, 2022, to September 15, 2022, at Euribor flat, that is, at no spread. If Euribor on June 13 turns out to be R, then that will be the rate applied to the loan, and the corporation will owe €10 million × $[1 + (92/360)R]$ on September 15. The corporation, therefore, is exposed to the risk that rates rise from January 14 to June 13.

The corporation can hedge this risk by locking in the forward rate of −0.47% as follows. On January 14, 2022, the corporation agrees to pay fixed on the FRA in the table. Then, on June 15, the corporation borrows

TABLE 12.9 A €10 Million Euribor Forward Rate Agreement from June 15, 2022, to September 15, 2022, at −0.47%, as of January 14, 2022.

Date	Description
1/14/2022	Trade Date
6/13/2022	3-Month Euribor Observed to be R
6/15/2022	Borrower Pays (net): $\dfrac{€10{,}000{,}000 \times \frac{92}{360}(-0.47\% - R)}{1 + \frac{92}{360}R}$

the €10 million it needs plus the quantity it owes on the FRA for three months at the then-prevailing rate, R. In total then, the corporation owes,

$$\left(€10{,}000{,}000 + \frac{€10{,}000{,}000 \times \frac{92}{360}(-0.47\% - R)}{1 + \frac{92}{360}R} \right) \left(1 + \frac{92}{360}R \right) \quad (12.13)$$

$$= €10{,}000{,}000 \left(1 + \frac{92}{360}R + \frac{92}{360}(-0.47\% - R) \right) \quad (12.14)$$

$$= €10{,}000{,}000 \left(1 + \frac{92}{360} \times (-0.47\%) \right) \quad (12.15)$$

on September 15, which is exactly the same as if it had borrowed €10 million at a rate of −0.47%. Intuitively, if realized Euribor on June 15 is less than −0.47%, the corporation borrows at relatively low rates, but owes money on the FRA. On the other hand, if realized Euribor is greater than −0.47%, the corporation borrows at relatively high rates but collects money on the FRA.[9]

An exchange-traded alternative to Euribor FRAs are Euribor futures contracts, some of which are listed in Table 12.10, as of January 14, 2022. The tickers begin with "ER," followed by the month and year indicators. The contracts are subject to daily settlement, and the final settlement rate is set to three-month Euribor on the last trade date. The final settlement rate of the June contract, for example, is three-month Euribor on June 13, 2022, which as mentioned earlier, represents the rate on a loan from June 15 to September 15. Three-month Euribor futures are scaled to hedge the change in interest on a €1 million, 90-day loan. Equivalently, a one-basis-point change in rate translates into a daily settlement flow of €1,000,000 × 0.01% × 90/360, or €25.

The corporation in the FRA example hedges a €10 million, 92-day loan from June 15 to September 15, 2022. This time period corresponds exactly to that covered by Euribor set on June 13, which is also the expiration of the June three-month Euribor contract. Furthermore, with the market at the levels in Table 12.10 – a rate of −0.47% – the corporation can sell the June contract to lock in a rate of −0.47% on its planned borrowing. The number of contracts required to hedge the €10 million, 92-day loan with a contract scaled to a €1 million, 90-day loan, is,

$$\frac{\$10{,}000{,}000 \times 92}{\$1{,}000{,}000 \times 90} = 10.22 \quad (12.16)$$

[9]The most creditworthy corporations might be able to borrow at Euribor flat, as in the example. Less creditworthy entities borrow at Euribor plus a spread, which, despite negative Euribor, might result in an overall positive interest payment. In any case, to the extent the bank can set the spread at the time of the loan, the corporation is exposed to spread risk.

TABLE 12.10 Selected Three-Month Euribor Futures Contracts, as of January 14, 2022. Rates Are in Percent.

Ticker	Delivery Month	Last Trade Date	Price	Rate
ERH2	Mar	03/14/22	100.540	–0.540
ERM2	Jun	06/13/22	100.470	–0.470
ERU2	Sep	09/19/22	100.375	–0.375
ERZ2	Dec	12/19/22	100.260	–0.260

TABLE 12.11 Hedging the Interest Cost of Borrowing €10 Million from June 15, 2022, to September 15, 2022, as of January 14, 2022, with June Three-Month Euribor Contracts, Priced Initially at –0.47%. The Hedge, Not Tailed, Sells 10.22 Contracts. Rates Are in Percent. Other Entries Are in Euro.

Final Settlement Rate	Loan Interest	Futures P&L	Net
–0.77	19,677.78	–7,665.00	12,012.78
–0.47	12,011.11	0.00	12,011.11
–0.17	4,344.44	7,665.00	12,009.44

The results of the hedge are given in Table 12.11. If rates rise, so that Euribor on June 13 turns out to be –0.17%, the corporation receives interest of €10 million × (92/360) × negative 0.17%, or €4,344.44; experiences a futures gain of 30 basis points per contract, for a total gain – ignoring the small effects of daily settlement – of 30 × €25 × 10.22, or €7,665.00; which together gives an overall gain of €12,009.44. Looking at the table as a whole, the futures position successfully hedges the cost of borrowing in the sense of locking in the receipt of about €12,011. The hedge would be exact if the number of contracts in Equation (12.16) were taken to more decimal places.

This hedging example is simple in that the corporation plans to borrow over the same period as covered by three-month Euribor relevant for the June contract. Along the same lines as the discussion in the context of hedging with SOFR futures, if the borrowing period were shorter or longer than that covered by the June Euribor contract, fewer or more contracts would be needed, and contracts other than June might be used.

The example in this section is constructed to show the similarity between Euribor FRAs and futures, but there are significant differences as well. FRAs may be subject to bilateral margin agreements, but are not subject to daily settlement, like futures. FRAs, as over-the-counter products, can be customized for individual trades with respect to dates covered and with

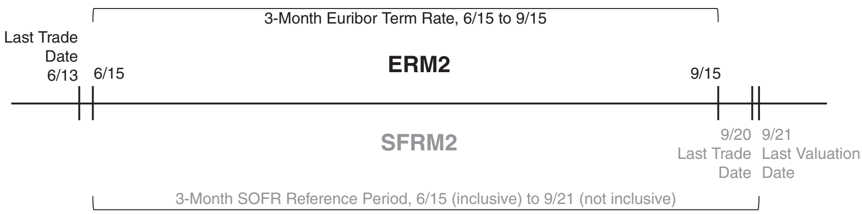

FIGURE 12.5 A Comparison of the June 2022 Euribor and the June 2022 SOFR Futures Contracts.

respect to notional amount, while the terms of futures contracts are highly standardized. In part because they are customized, however, FRAs are less liquid than futures. Hedgers typically need to decide, therefore, whether customization or liquidity is more important in their particular circumstances.

This section concludes with a comparison of futures on a forward-looking term rate benchmark, like Euribor, and futures on a backward-looking compounded overnight benchmark, like SOFR. The top half of Figure 12.5 depicts the June three-month Euribor contract, ERM2, while the bottom half depicts the June three-month SOFR contract, SFRM2. Both contracts cover realized interest rates from mid-June to mid-September, though the terminal date differs somewhat, as detailed earlier. The big difference, however, is that ERM2, which references a term rate set on June 13, expires on that date. SFRM2, which references a compounded daily rate that is not known until September 21, does not expire until that later date.

12.6 THE FUTURES-FORWARD DIFFERENCE

By definition, borrowers and lenders can lock in market forward rates, and earlier sections of this chapter show that, ignoring daily settlement, borrowers and lenders can also lock in market futures rates. This section explains that, because of daily settlement, market futures rates exceed market forward rates. However, the magnitude of the difference turns out to be small except for contracts of the longest terms.

If a trader receives fixed at 2% through a forward agreement, all of the P&L from the agreement is realized at its expiration. But if the trader buys an otherwise identical futures contract at 2%, that same total P&L is realized over time. More specifically, when rates are falling and the futures contract is making money, daily settlement profits are reinvested at relatively low rates. And when rates are rising and the futures contract is losing money, daily settlement losses have to be financed at relatively high rates. Therefore, averaging across scenarios of both falling and rising rates, receiving daily settlement payments over time is worse than receiving all of the

P&L at the end. Hence, traders require a higher rate when buying a futures contract than when receiving fixed in a forward agreement. Or, in reverse, other traders willingly accept a higher rate when selling a futures contract than when paying fixed in a forward agreement. In this example, the futures rate exceeds the forward rate of 2%. Furthermore, based on the logic of that result, the difference between the two rates increases with the length of time to expiration of the contracts and with the volatility of interest rates.

Appendix A12 proves, in general, that the futures rate exceeds the forward rate.[10] For the purposes of understanding orders of magnitude, however, Table 12.12 presents some results, without proof, from a normal one-factor term structure model with a constant drift and with a constant volatility of 80 basis points per year. This volatility is chosen to be roughly consistent with the levels shown in Chapter 16. In any case, the rows of the table in the first and second panels correspond to the three different kinds of futures contracts described in this chapter: three-month contracts on a term rate (Euribor futures); one-month contracts on an average of overnight rates (fed funds and SOFR futures); and three-month contracts on compounded overnight rates. The first panel gives the formulas for the differences between futures and forward rates in the model just described, where σ denotes the annual normal volatility (e.g., 0.8% for 80 basis points); β denotes the term of the underlying rate, in years, which, for the contracts

TABLE 12.12 The Futures-Forward Rate Difference for Selected Contracts in a One-Factor Model with a Volatility of 80 Basis Points per Year.

Underlying Rate	Contract Examples	Formula		β	
Term	3m Euribor	$\frac{\sigma^2 t^2}{2} + \frac{\sigma^2 t\beta}{2}$		0.25	
Avg. ON	1m FF/SOFR	$\frac{\sigma^2 t^2}{2} + \frac{\sigma^2 t\beta}{2} + \frac{\sigma^2 \beta^2}{6}$		$\frac{1}{12}$	
Comp. ON	3m SOFR	$\frac{\sigma^2 t^2}{2} + \frac{\sigma^2 t\beta}{2} + \frac{\sigma^2 \beta^2}{3}$		0.25	
			Fut-Fwd Difference (bps)		
		$t = 1$	$t = 2$	$t = 5$	$t = 10$
Term	3m Euribor	0.4	1.4	8.4	32.8
Avg. ON	1m FF/SOFR	0.3	1.3	8.1	32.3
Comp. ON	3m SOFR	0.4	1.5	8.4	32.8

[10]This result is analogous to the result in Chapter 11 that the futures price is less than the forward price.

shown, is either 0.25 or 1/12; and t denotes the time to the beginning of the rate reference period. As an illustration of t from the examples in the text, the beginning of the reference period of the three-month June Euribor and three-month June SOFR contracts depicted in Figure 12.5 – June 15 – is 152 days or about $t = 0.42$ years from the trade date of January 14.

The second panel of Table 12.12 gives the differences between the futures and forward rates in the model for each of the contract types and for four values of t, in basis points. The first point to notice is that futures-forward rate differences for a given t do not vary much across contracts. The reason for this can be seen from the formulas in the first panel. With β relatively small, the dominant term in all the formulas is the common term $\sigma^2 t^2/2$. The second point to notice is that the futures-forward rate difference is relatively small except for the longest times to the beginning of the reference period. This is again explained by the dominant term, $\sigma^2 t^2/2$, which increases with the square of t. That term alone, in fact, is a good approximation for all of the terms in the given formulas: for $t = 10$, for example, $\sigma^2 t^2/2$ equals 32 basis points, which approximately equals the exact solution for all of the contract types.

CHAPTER 13

Interest Rate Swaps

Interest rate swaps (IRS) are contracts in which two counterparties agree to exchange a sequence of interest payments on some *notional amount* of currency. In an *overnight index swap (OIS)*, payments based on a fixed rate of interest are exchanged for payments based on a *floating*, overnight rate, which changes daily with market conditions. The swaps introduced in Chapter 2 are OIS, in which the floating rate is the *Secured Overnight Financing Rate (SOFR)* defined and discussed in Chapters 10 and 12. A *fixed-for-floating* swap is similar, but the floating rate is a term rate, rather than an overnight rate. Euribor swaps are fixed-for-floating swaps, in which the floating rate is typically the three-month Euribor rate described in Chapter 12. Historically, the most common IRS across currencies were fixed-for-floating *London Interbank Offered Rate (LIBOR)* swaps, in which the floating rate was LIBOR of some term. However, with the transition away from LIBOR, as described in Chapter 12, these swaps are disappearing. OIS and fixed-for-floating swaps are the main focus of this chapter.

There are several classes of derivatives closely related to IRS, including *forward-rate agreements (FRAs)* (see Chapter 12), caps and floors, and swaptions (see Chapter 16). While traditionally not called "swaps" by the financial industry, these products are defined as "swaps" in the Dodd-Frank Act, which can cause some terminological confusion. In any case, these products are discussed elsewhere in this book but are included in the market size statistics presented in this chapter.

The first section of this chapter describes the size of the IRS market and how various market sectors use swaps. The second section builds on the introduction of IRS in Chapter 2 and of short-term rates in Chapter 12 to present more detail on cash flows, pricing, and risk metrics. The third section uses several examples and cases to illustrate how IRS are used to manage risk. The fourth section addresses the risk that a swap counterparty defaults on contractual obligations, that is, *counterparty credit risk*, and describes the posting of *collateral* or *margin* to mitigate this risk. The fifth section explains swaps *clearing*, through which the two counterparties to a swap legally face a *clearinghouse* instead of each other. The sixth section introduces *basis swaps* and *basis swap spreads* and explains how swaps that reference nearly risk-free, floating-rate indexes, like SOFR, are priced differently from swaps referencing other floating-rate indexes.

13.1 MARKET SIZE AND PARTICIPANTS

Table 13.1 shows outstanding amounts of interest rate swaps.[1] Products include OIS, fixed-for-floating swaps, FRAs, interest rate caps and floors, and swaptions. Counterparties include all those that report positions to the US Commodity Futures Trading Commission (CFTC), that is, US entities, US subsidiaries of foreign entities, and foreign swap dealers that register with the CFTC so as to do business with US persons. The first column of the table divides these reporting entities into selected sectors. The second and third columns give the total long and short notional amounts in each sector. As explained in Chapter 2, the notional amount of a swap is the amount on which payments are based. And, to be consistent with bond terminology, "long" positions are those that increase in value when interest rates fall (e.g., receiving fixed), while "short" positions are those that increase in value when interest rates rise (e.g., paying fixed). Because every swap contract has one counterparty long and the other short, total longs must equal total shorts across the whole market. According to Table 13.1, the total notional amount of the market is $210.7 trillion. This is an enormous number compared to the sizes of markets as presented in Chapter O: the total amount of debt and loans outstanding in the United States in Figure O.4 is only $76 trillion. As it turns out, however, notional amount outstanding vastly exaggerates the size of the IRS market.

The fourth and fifth columns of the table give a long and short "five-year equivalent" notional amount, which is the notional amount of five-year swaps that has the same risk as the raw notional amount. The risk sensitivity of swaps is discussed in the next section, but say, for example, that one sector was long $100 million of 10-year swaps, which have a DV01 of 0.090; that is, a one-basis-point decline in rates increases the value of the position by $100 million times 0.090/100, or $90,000. Say also that five-year swaps have a DV01 of 0.045. Then, the actual swap position – $100 million of 10-year swaps – has the same risk as $200 million of five-year swaps, and its "five-year equivalent" notional amount is $200 million. On the other hand, a position of $100 million two-year swaps, which has a DV01 of 0.018, has

[1]The discussion of this section is motivated by Baker, L., et al. (2021), "Risk Transfer with Interest Rate Swaps," *Financial Markets, Institutions & Markets* 30(1), New York University Salomon Center and Wiley Periodicals.

TABLE 13.1 Entity-Netted Notionals (ENNs) of Interest Rate Swaps, US Reporting Entities, as of September 2020, in $Trillions.

	Notionals		5-Year Equivalents		ENNs		ENNs
(1)	Long (2)	Short (3)	Long (4)	Short (5)	Long (6)	Short (7)	Net (8)
Swap Dealer	158.7	158.8	100.6	98.9	9.3	7.6	−1.7
Hedge Fund	19.5	16.8	6.3	6.2	0.9	0.7	−0.1
Bank	18.2	19.5	15.2	17.2	1.3	3.2	−1.9
Other Financial	6.3	6.2	5.4	5.2	0.8	0.6	−0.2
Asset Manager	2.6	2.8	1.9	2.2	0.7	1.0	−0.2
Pension	1.6	1.2	3.0	2.1	1.2	0.3	−0.9
Gov't/Quasi-Gov't	1.6	1.7	1.4	1.7	0.4	0.7	−0.3
Nonfinancial	0.3	1.4	0.2	1.0	0.2	1.0	−0.8
Insurance	1.1	0.9	2.4	1.7	1.0	0.3	−0.7
Other	0.9	1.4	0.8	1.1	0.5	0.8	−0.4
Total	210.7	210.7	137.1	137.1	16.1	16.1	−0.0

Products include fixed-for-floating swaps, forward-rate agreements, overnight index swaps, swaptions, and interest rate caps and floors. *Source:* Baker, L., and Mixon, S. (2020), "Introducing ENNs: A Measure of the Size of Interest Rate Swap Markets," Update as of September 11, Office of the Chief Economics, Commodity Futures Trading Commission, September.

a five-year equivalent notional amount of $100 million times 0.018/0.045, or $40 million. Returning to Table 13.1, the total five-year equivalents in the market are $137.1 trillion, which is significantly less than the total notional amount of $210.7 trillion. First, the options on swaps included here have lower risk than the risk of their underlying notional amounts.[2] Second, a large proportion of swaps have terms less than five years. In fact, because bond markets, like the US Treasury and corporate bonds markets, have maturities concentrated in the five- to 10-year range, five-year swap equivalents are better than raw notional amounts for comparing market sizes.

The sixth and seventh columns of Table 13.1 report *entity-netted notionals (ENNs)*, which net long and short five-year equivalents that are denominated in the same currency between each pair of counterparties. For example, if Counterparty A is receiving fixed from Counterparty B on $100 million five-year equivalents and simultaneously paying fixed to Counterparty B on $60 million five-year equivalents, then Counterparty

[2]The interest rate sensitivity of an option on a swap is the risk of the notional amount of the underlying swap times the option's delta, which is always less than one except for some exotic products, which are not included in this analysis.

A is long \$40 million ENNs against Counterparty B and, symmetrically, Counterparty B is short \$40 million ENNs against Counterparty A. Summing long or short ENNs across all counterparty pairs in a sector or across the market as a whole gives the corresponding elements of the table. Total market ENNs are \$16.1 trillion, which are much less than five-year equivalents. As explained in the next section, participants in the IRS market often take off risk by taking on new, risk-offsetting positions, not by unwinding existing trades. As a result, pairs of counterparties are often both long and short to each other, and their ENNs are much less than their five-year equivalents. In any case, at \$16.1 trillion, the size of the IRS market is comparable to that of other US fixed income markets, as reported in Chapter O.

Swap dealers, whose business is to make markets in swaps, are very likely to accumulate both long and short positions with their clients and are even more likely to do so with swaps clearinghouses, which are the legal counterparties to all cleared trades. In fact, the reduction of dealers' approximately \$100 trillion long and short five-year equivalents to \$9.3 trillion long and \$7.6 trillion short ENNs accounts for most of the reduction of total five-year equivalents to total ENNs. Significant reductions also characterize hedge funds, which are in the business of trading swaps, and banks, which, as described later, use hedged swap positions to facilitate their customer loan businesses.

The final column of Table 13.1 is just the difference between long and short ENNs, which reveals whether a sector, as whole, is long or short swaps. (Differences are not exact because of rounding.) There is probably too much variation across hedge funds and asset managers, with each entity pursuing its own investment strategies, to explain the signs of these sectors' net ENNs. Banks as a sector are short, perhaps as a result of hedging the interest rate risk of their mortgage assets, although, as discussed next, banks and other financial companies receive fixed to transform their long-term fixed-rate debt into synthetic floating-rate debt. Pension funds and insurance companies are likely overall long, as discussed below, to hedge the interest rate risk of their long-term liabilities. Finally, also as discussed presently, nonfinancial companies pay fixed both to transform their floating-rate bank loans into synthetic fixed-rate loans and to hedge the risk of increasing rates before anticipated sales of long-term debt.

13.2 IRS CASH FLOWS AND ANALYTICS

Chapter 2 describes the cash flows and pricing of a SOFR swap, and Chapter 12 explains the difference between borrowing or lending at rolling overnight rates *versus* term rates. This section builds on those foundations.

Cash Flows of OIS versus Fixed-for-Floating Swaps

Figure 13.1 illustrates the cash flows of a \$100 million three-year 1.64% versus SOFR swap, which is an OIS, while Figure 13.2 illustrates the cash flows of a €100 million three-year 0.36% versus six-month Euribor swap, which is a fixed-for-floating swap. Both swaps settle on February 16, 2022, and both figures are from the perspective of the fixed receiver. The SOFR swap makes annual cash flows using the actual/360 convention on both its fixed and floating legs. Therefore, the fixed interest earned after each 365-day year, from February 16, 2022, to February 16, 2023, and from February 16, 2023, to February 16, 2024, is $\$100{,}000{,}000 \times 1.64\% \times 365/360 = \$1{,}662{,}778$. Similarly, the fixed interest earned after the 366-day year from February 16, 2024, to February 16, 2025, is $\$100{,}000{,}000 \times 1.64\% \times 366/360 = \$1{,}667{,}333$.[3] The floating cash flow of the SOFR swap over a given year, as explained in Chapter 2, equals the daily-compounded interest

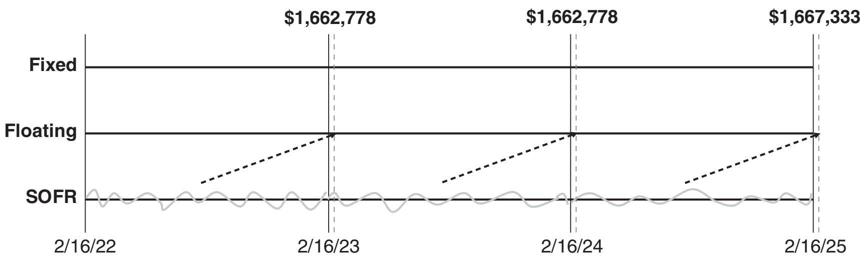

FIGURE 13.1 Receiving Fixed at 1.64% on a \$100 Million Three-Year SOFR Swap.

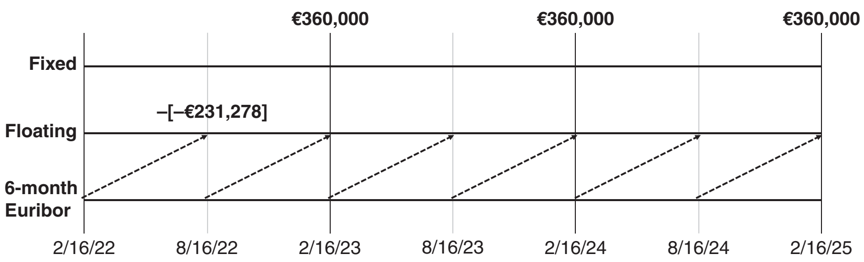

FIGURE 13.2 Receiving Fixed at 0.36% on a €100 Million Three-Year Fixed versus Six-Month Euribor Swap. Initial Euribor Setting Is −0.46%.

[3]This presentation somewhat simplifies payment conventions. For example, because February 16, 2025, is a Sunday, the payments scheduled for that day are actually paid on the following business day, February 17, 2025, with additional interest for that extra day.

earned on $100 million at realized SOFR rates over that year. Because each year on this swap starts and ends on February 16, the relevant SOFR rates are those from and including February 16 of one year through February 15 of the next year. Furthermore, because SOFR on February 15 is not published until the morning of February 16 (see Chapter 12), swap counterparties are given until February 17 to make payments on both legs of the swap. This *payment delay* is indicated in the figure by the light vertical lines to the right of February 16 of each year and by floating payment arrows pointing slightly past February 16.

Fixed-for-floating swaps use realized term rates to determine floating-rate payments. The payment frequency on the fixed leg of these swaps can be annual or semiannual, while the payment frequency on the floating leg usually matches the tenor of the index rate. Figure 13.2 illustrates the cash flows of a swap of a fixed rate of 0.36% against six-month Euribor. The fixed side of this swap pays annually and follows the 30/360 convention. Recalling that there are always 360 "30/360 days" in a year, the annual payment on the fixed leg is simply €100 million × 0.36% × 360/360, or €360,000. The floating side of the swap follows the actual/360 convention and pays semiannually to match the tenor of the floating-rate index, namely, six-month Euribor. The rate used to determine each floating-rate payment is the Euribor setting two business days before the beginning of the payment period.[4] Because six-month Euribor on February 14, 2022, was −0.46%, and because there are 181 days between February 16 and August 16, 2022, the first floating-rate payment of the swap in Figure 13.2 is €100,000,000 × (−0.46%) × 181/360, or −€231,278. This means that the fixed receiver actually receives €231,278 from the floating leg of the swap. Note that there is no payment delay in fixed-for-floating swaps, because the cash flows are known well in advance of each payment date.

Swap Valuation: More on the Floating Leg

Chapter 2 explains the trick in valuing SOFR swaps, that is, adding a fictional payment of the notional amount at maturity to both legs of the swap. This trick is so much a part of swap valuation that the terms "fixed leg" and "floating leg" typically include these fictional notional amounts. In any case, the fixed leg of a swap now looks like a coupon bond, which makes periodic interest payments and a final "principal" payment and can be priced with the tools of the early chapters of this book. The floating leg of a swap now looks like a floating-rate bond. In terms of valuation, this means that the floating leg is worth par or notional amount on all reset dates.

[4]Euribor is meant to represent the rate on a Euro deposit settling in two days. Therefore, the appropriate floating rate for any six-month period is the six-month Euribor published two days before the beginning of that period.

To elaborate on this last point, in the example of Figure 13.1, the floating leg is worth \$100 million on each February 16 of 2022, 2023, and 2024. Intuitively, because SOFR is the fair, risk-free market rate of interest, investors are willing to pay exactly \$100 million at the beginning of each annual period for a "floating-rate bond" that pays compounded SOFR interest at the end of each year. Between reset dates, the value of the floating leg of the swap equals the accrued value of its notional amount at the already realized SOFR rates. Say, for example, that SOFR over the first 100 days of the swap was 0.05%, giving an accrued value of $\$100{,}000{,}000 \times (1 + 0.05\%/360)^{100} = \$100{,}013{,}890$. From that point on, the floating leg continues to accrue at the fair market rate, which is SOFR, and will be worth \$100 million, after interest is paid out, on February 16, 2023. Hence, after those 100 days, investors are willing to pay exactly \$100,013,890 for the hypothetical floating-rate bond. To summarize, inclusive of the fictional notional amount, the floating leg of the SOFR swap is like a deposit that pays the fair market rate of interest. It accrues value until all interest earned is withdrawn (i.e., on reset dates), at which point its value falls back to the original deposited amount.

Valuing the floating leg of the Euribor swap is somewhat different. Unlike OIS, fixed-for-floating swaps determine floating-rate payment amounts at the beginning of each period. This does not affect valuations on reset dates – with an important caveat given next – but does change valuations between reset dates. To explain, turn again to the swap in Figure 13.2. The cash flow of the floating leg of the swap on February 16, 2025, is €100 million plus interest at six-month Euribor effective from August 16, 2024. But if Euribor is appropriate for discounting swap cash flows, then, by definition, the present value of that cash flow as of August 16, 2024, is €100 million. Now step back to February 16, 2024. On August 16, 2024, the floating leg is worth €100 million, as just shown, plus interest at six-month Euribor effective from February 16, 2024. Once again, then, by definition, the present value of the floating leg as of February 16, 2024, is €100 million. Proceeding backward along these lines shows that the floating leg is worth €100 million on each reset date. Between reset dates, however, the value of the floating leg is the discounted value of €100 million plus the interest payment already determined as of the previous reset date. In other words, the value of the floating leg between reset dates equals the value of a zero coupon bond that matures on the next reset date, paying par plus the interest set on the previous reset date. Another difference between pricing SOFR and six-month Euribor swaps, however, is that six-month Euribor is not a risk-free rate: as an interbank term rate, it includes a spread due to the credit risk of the banking sector. Therefore, discount factors and rates derived from Euribor swaps may not be appropriate for valuing the cash flows of these swaps. This point is ignored for the present but revisited in the last section of the chapter.

Net Present Value (NPV)

At initiation of a swap, its value is zero. The fixed rate in the market at any moment, in fact, is such that the value of the fixed and floating legs are equal or, equivalently, such that the initiation of the swap requires no exchange of money upfront. As time passes, however, a swap increases in value to one counterparty and decreases to another. To illustrate, return to the three-year SOFR swap in Figure 13.1. At initiation, it has a fixed rate of 1.64%, which sets the present value of both its fixed and floating legs to $100 million. Say that, one year later, just after the first interest payments on each leg of the swap, the fair two-year SOFR swap rate is 1%. Because rates have declined, the present value of the remaining cash flows of the fixed leg of the 1.64% swap increases, say to $101,280,000. The value of the floating leg of the swap, on a reset date, is $100,000,000. Therefore, the NPV of the swap to the fixed receiver is $101,280,000 minus $100,000,000, or $1,280,000. The fixed receiver of 1.64% is in-the-money on the now two-year swap, because traders initiating new two-year swaps are receiving only 1%. Conversely, the NPV of the swap to the fixed payer is −$1,280,000: the fixed payer of 1.64% is out-of-the-money, because new traders are paying only 1%.

With NPV defined, the text can mention *gross market value*, which is defined as the sum of the absolute values or positive NPVs across swaps in a given market. Gross market value is sometimes used as a measure of exposure in a market but suffers from two weaknesses along these lines. First, positive and negative NPVs within counterparty pairs are added in absolute value, rather than netted. Second, to the extent that margin is posted against NPV, gross market value can significantly overstate counterparty exposure.

DV01 of a Swap

With the fixed leg conceptualized as a coupon bond, its DV01 can be computed along the lines of Chapter 4, that is, by shifting the appropriate rate curve down by one basis point and revaluing. Because the value of the floating leg of a SOFR swap at any time equals its accrued value, no matter how rates might move at that moment, its DV01 is zero. And because the value of the floating leg of a Euribor swap is equal to that of a bond paying interest and “principal” and maturing at the next reset date, its DV01 can also be computed along the lines of Chapter 4. To clarify with some orders of magnitude, at a flat term structure of 2%, the DV01 of the fixed leg of a 2% 10-year swap is like that of a 10-year bond, or about 0.09; the DV01 of the floating leg of the SOFR swap is zero; and the DV01 of the floating leg of the fixed versus six-month Euribor swap is like that of a zero coupon bond with six months to maturity, or about 0.005. In the case of the Euribor swap, it could be said that the DV01 of receiving fixed is the DV01 of the fixed leg minus the DV01 of the floating leg, or, in the numerical example, 0.09 – 0.005 = 0.085. As argued in Chapter 5, however, bonds subject to different parts of the term structure are far from perfectly correlated. Therefore, the 10-year fixed leg and the essentially six-month floating leg should be hedged separately, the former with other instruments having about 10 years to maturity, and the latter with other instruments having about six months to maturity. Recall, by the way, that Section 5.5 discusses *PV01* and *partial '01s*, which specifically address the interest rate sensitivities of swap books.

Unwinding IRS Risk Positions

Say that the counterparties of the three-year SOFR swap in Figure 13.1 are Counterparty A, who is receiving fixed, and Counterparty B, who is paying fixed. Say further that Counterparty A decides to take off the interest rate risk of the swap one year later, in the scenario described already, in which the prevailing two-year SOFR swap rate is 1% and the NPV to Counterparty A of the existing swap is $1,280,000. There are typically three ways in which Counterparty A can proceed. First, Counterparty A can ask Counterparty B to unwind the trade. Because the NPV is positive to Counterparty A, Counterparty B would pay Counterparty A that NPV and the two would then tear up the trade. An advantage of this approach is that the trade really disappears, just like a particular bond or futures contract that is bought and subsequently sold. Unfortunately for Counterparty A, however, Counterparty B might want to keep the trade alive: this swap might be part of a larger trading portfolio; paying the NPV immediately might entail funding costs; and the trade's removal might increase overall counterparty risk exposure to Counterparty A. Also, because only Counterparty B can unwind this particular trade, Counterparty B has negotiating leverage and might offer to pay less than the theoretically fair NPV. In any case, in practice, counterparties like Counterparty A find it difficult to unwind existing trades.

A second way for Counterparty A to take off the interest rate risk of the swap with Counterparty B is to pay 1.64% for two years to another counterparty, receiving at initiation $1,280,000 as the NPV of paying 1.64% in a 1% rate environment. This trade would flatten the interest rate risk of the existing swap: paying 1.64% to and receiving SOFR from this new counterparty exactly offsets receiving 1.64% from and paying SOFR to Counterparty B. This approach has problems too, however. It may be difficult to find a counterparty who wants to receive 1.64% on a two-year swap and is willing to pay some NPV to do so, when the market rate is 1%. Put another way, the most liquid two-year swap contracts have a fixed rate equal to the prevailing market rate and are worth zero at initiation. Another problem with the approach is that the original swap does not disappear. Over the next two years, both the swap with Counterparty B and the swap with the new counterparty coexist, and Counterparty A has to manage or bear the counterparty risk on both.

The third way for Counterparty A to take off the interest rate risk of the swap, which is the most common in practice, is to pay 1.0% on a new two-year swap, but adjust its notional amount so as to hedge the risk of the existing swap. In this example, with these two-year swaps differing only in the fixed rates, and those only by 0.64%, the adjustment will be extremely small. To illustrate the general point, therefore, say that Counterparty A was offsetting the risk of receiving fixed on an existing $100 million 25-year swap, which has a DV01 of 0.196, by paying fixed on a new 30-year swap, which has a DV01 of 0.225. In that case, Counterparty A would pay fixed on $100 million times 0.196/0.225, or on about $87.1 million of the 30-year swap. The advantage of this approach is that, by choosing to trade the most liquid swap, Counterparty A can relatively easily obtain a competitive rate. The problems with the approach are that the two swaps are not quite risk offsetting, that is, the hedge might have to be adjusted over time, and that, as in the second approach, both the original swap and the new swap remain in existence with their respective counterparty risks.

The practice of taking off swap interest rate risk not by unwinding trades, but by initiating offsetting trades, gives rise to notional amounts that exaggerate interest rate risk, as described in the previous section, and to a heightening of operational complexities in the IRS market. Furthermore, to the extent that net long positions are held against different counterparties than net short positions, counterparty risk must be managed or borne on both the long and short sides. The industry does mitigate the proliferation of trades by regular *compression* programs, which cancel trades across the system, accompanied by relatively minor cash payments, in a way that minimally changes the risk profiles of participating counterparties. The predominance of IRS clearing mitigates the counterparty risk implications of the proliferation of trades, as trades are transformed so that each counterparty faces the clearinghouse as its legal counterparty.

A fundamental question is why the industry has not migrated to practices that better facilitate the unwinding of swaps. As explained in Chapter 14, the proliferation of credit default swaps was significantly reduced by meaningful standardization of contract coupon and maturity. Similarly standardized IRS do exist, in the form of *Market Agreed Coupon (MAC)* swaps, but volumes are low. Standardized swap futures are another possibility, and they do exist, but also trade in relatively limited volumes. One explanation for why the IRS market has not moved in these directions is that market participants actively want specific coupons and maturities. And there are hedging applications – some discussed herein – in which customized swaps are preferable. A recent study, however, argued that between about 60% and 80% of IRS trading volume can be considered standardized, which implies that significant volumes could be migrated to practices that reduce trade proliferation.[5]

13.3 USES OF INTEREST RATE SWAPS

IRS can be conceptualized in two ways: as a leveraged position in a bond (see Chapter 2) and as an exchange of interest payments. The first interpretation best explains the actions of market participants in increasing or decreasing their interest rate exposures. An asset manager deciding to pay fixed, either as a bet that interest rates will rise or as a hedge against the interest rate risk of a corporate bond portfolio, is more interested in the DV01 of the swap than in the particular cash flows being exchanged. The first three use cases described in this section – pensions, Greece, and Brunswick Corporation – fall into this category. The second interpretation of swaps best explains the actions of market participants wanting to alter particular streams of cash flows. In the asset swap transactions described in Chapter 14, the owner of a corporate bond specifically wants to substitute the receipt of bond coupon with the receipt of a floating rate plus a spread. The last two use cases described here – bank loans and synthetic floating-rate debt – fall into this category.

Pension Liabilities

Say that a pension fund has long-term liabilities with a present value of \$1 billion and a DV01 of \$2 million. At the same time, based on its research across corporate credits, it has invested \$1 billion in a portfolio of corporate bonds. However, because the universe of corporate bonds lies mostly in the five- to 10-year maturity range, and because the process of choosing credits does not focus on maturity, the selected portfolio has a DV01 of only \$500,000. This configuration exposes the pension fund to the risk that interest rates fall: for every one-basis-point decline in rates, the value of the liabilities increase by \$2 million and the value of the assets increase by \$500,000, causing a funding gap of \$1.5 million. To hedge, the pension fund can receive fixed in a portfolio of long-term IRS with a total DV01 of \$1.5 million. The exact distribution of swaps across maturities is likely to be chosen to hedge against changes in different parts of the term structure of rates, as explained in Chapter 5. In any case, note that receiving fixed

[5]Haynes, R., Lau, M., and Tuckman, B. (2020), "How Customized Are Interest Rate Swaps?" Office of the Chief Economist, Commodity Futures Trading Commission, June.

in swaps does not require cash from the pension fund, except for posting margin, as described presently. In other words, the pension fund uses its cash to buy the assets it really wants, in this example corporate bonds, and then adjusts the resulting risk profile by trading swaps.

Greece Hedges Floating-Rate Debt

During the sovereign debt crisis in 2010, Greece had borrowed money from other Eurozone countries at Euribor plus a spread.[6] Toward the end of 2018, the loan balance was over €50 billion, and 10-year swap rates had fallen to around 1.60%. Greece chose, therefore, to hedge the risk of future increases in Euribor by paying fixed in 10-year Euribor swaps. The problem, however, was that Greece had a below investment-grade credit rating and, as a sovereign, did not post collateral. This meant that Greece was not a particularly attractive swap counterparty, particularly to dealers who had positive NPVs on existing swaps with Greece. More specifically, i) without collateral, the positive NPVs of dealers translated into outright credit exposures to Greece; and ii) these dealers had negative NPV on the hedges of their IRS with Greece, against which they typically had to post collateral and incur funding costs.[7] To overcome the reluctance of these dealers, Greece allowed them to *novate* their positive NPV swaps to dealers with negative NPV swaps. For example, if Dealer X had a swap with Greece with an NPV of €1 million, and Dealer Y had a swap with Greece with an NPV of −€1 million, then Dealer Y could pay Dealer X €1 million and take over its swap. Both dealers would then have no current exposure to Greece; Dealer X could unwind its hedge and its associated funding cost; and, consequently, Dealer X would no longer be reluctant to trade new swaps with Greece.[8]

Hedging Future Debt Issuance

This use case starts by describing a *forward swap* or a *forward-starting swap*, with Figure 13.3 illustrating a three-year swap, one year forward. Through this forward swap, one counterparty agrees to pay fixed and receive floating for three years starting in one year, while the other agrees to receive fixed and pay floating. Importantly, the fixed rate is agreed on as of the trade date.

---

[6]The facts of this case can be found in Becker, L. (2018), "Greece Slashes Rates Exposure with €35 Billion Swap Programme," *Risk.net*, November 28.

[7]This is a typical example of why a *funding value adjustment (FVA)* is needed to evaluate swap positions. The net costs or benefits of posting and receiving collateral must be considered alongside the cash flows of the swap.

[8]Dealer Y, representative of those with negative NPV to Greece, might require additional incentives to do this novation. It might not mind owing money to Greece, and, as it likely has positive NPV on its hedges against which it has collected collateral, might be saving on overall funding costs.

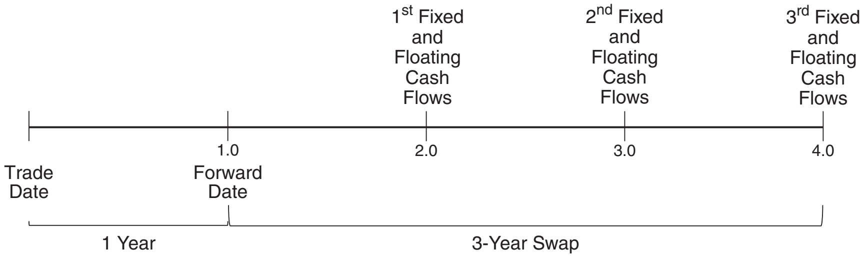

FIGURE 13.3 A Three-Year Swap, One-Year Forward, with Annual Interest Payments.

Now consider a corporation that plans to sell $100 million of three-year bonds, but in one year's time. Assume that the corporation will be able to sell these bonds at a rate 1.5% over the then-prevailing three-year swap rate, and that the current three-year swap rate, one year forward is 1.0%. The corporation can then lock in a borrowing cost of 1% + 1.5% = 2.5% by paying fixed today on $100 million of the 1.0% three-year swap, one year forward. Table 13.2 shows how the hedge works. Say that, one year later, the three-year swap rate has risen to 1.5%. The corporation then issues $100 million of three-year bonds at 3.0% and receives fixed at 1.5% on a $100 million three-year swap. The net interest from the existing 1.0% forward swap, which has become a three-year swap, the debt issue, and the new swap is –1.0% – 3.0% + 1.5% = –2.5%. Alternatively, say that, one year later, the three-year swap rate has fallen to 0.5%. The corporation then issues at 2.0% and receives in a three-year swap at 0.5% for the same net of –1.0% – 2.0% + 0.5% = –2.5%. Essentially, if rates rise, the corporation pays more on its debt, but wins on having locked in payment of a now low rate on the forward-starting swap. Conversely, if rates fall, the corporation pays less on its debt, but loses on having locked in payment of a now high rate on the forward-starting swap.

TABLE 13.2 Debt Is Sold at 1.5% over the Swap Rate. Hedge Future Debt Issuance by Paying Fixed at 1% on a Three-Year Swap, One Year Forward. The Three-Year Swap Rate in One Year Is Either 1.5% or 0.5%.

3-Year Swap Rate	1.5%	0.5%
Trade	**Cash Flow**	**Cash Flow**
Issue Debt	–3.0%	–2.0%
Receive Fixed	1.5%	0.5%
Pay Fixed on Fwd Starting Swap	–1.0%	–1.0%
Total	–2.5%	–2.5%

In practice, instead of receiving fixed on a new swap at the time of issue, the corporation can unwind the forward-starting swap and apply its realized NPV to the debt issue. A positive NPV would reduce the higher cost of debt, perhaps by using that positive NPV as proceeds and selling less debt, while a negative NPV would increase the lower cost of debt, perhaps by selling more debt to pay off that negative NPV.

Debt issuers sometimes find that forward swaps are expensive to transact relative to spot-starting swaps. In that case, an issuer might pay fixed in a spot-starting swap on a notional amount set so the DV01 of the hedging swap matches the DV01 of the forward sale of corporate bonds. Hedges with spot-starting swaps do entail some curve risk, however. For example, in the scenario of the previous paragraph, paying fixed on a spot-starting three-year swap leaves the debt, but not the hedge, exposed to changes in the one-year rate, three years forward. And paying fixed on a spot-starting four-year swap leaves the hedge, but not the debt, exposed to changes in the one-year rate. In any case, note that swap hedges protect only against changes in the swap rate. Any increases in corporate borrowing rates relative to the swap rate, which was assumed away in the example of the previous paragraphs, are clearly not hedged by swap positions.

The Brunswick Corporation, which manufactures boats and other products, planned to issue $300 million of fixed-rate debt in mid-2023. In January 2021, it entered into $150 million of forward-starting swaps to hedge half of that issuance. Furthermore, it seems to have structured its hedges with various starting dates and maturities to account for uncertainty about the actual issuance dates. The treasurer indicated that the volume of hedges might increase over time as, presumably, the dates of issuance become firmer. In short, figuring out what needs to be hedged can be much more difficult than figuring out how to hedge.[9]

Bank Loans

Banks prefer to make floating-rate loans to customers. First, from the perspective of asset–liability management, floating-rate assets naturally hedge deposits, which are floating-rate liabilities (and which constitute the overwhelming majority of bank funding). Second, as discussed in Chapter 14, floating-rate loans are much easier to sell in the secondary market than are fixed-rate loans. Many borrowers, however, want to lock in a fixed interest rate over some term of their borrowings. The reconciliation of bank and borrower objectives is achieved with IRS, as illustrated in Figure 13.4. The

[9]The description of Brunswick's hedging program is from Turnstead, R. (2021), "Corporates Pre-hedge Future Bond Sales as Inflation Rises," *Risk.net*, August 17.

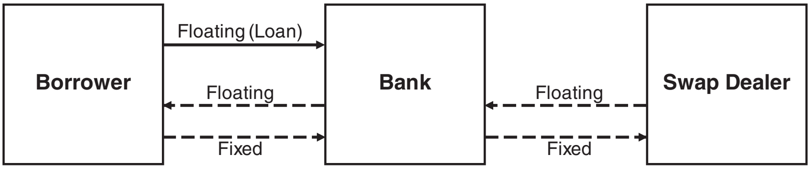

FIGURE 13.4 Facilitating a Bank Loan with an Interest Rate Swap.

bank makes a floating-rate loan to the borrower, on which the borrower pays a floating rate. At the same time, the bank receives fixed from the borrower and, in a *back-to-back* swap, pays fixed to a dealer. The net result for the customer is a fixed-rate loan: paying floating on the loan is offset by receiving floating on the swap, leaving only the fixed payments on the swap. The net result to the bank is the original floating-rate loan: the cash flows from the two swaps, by design, exactly offset each other. While many banks do facilitate their customer loan business as illustrated in the figure, with back-to-back swaps, some hedge the aggregate risk of their customer swap portfolio with relatively few larger swaps.

Synthetic Floating-Rate Debt

As just discussed, banks borrow at floating rates through deposits and make floating-rate loans. Banks usually do want some amount of long-term funding, however, because an overreliance on short-term funding creates a vulnerability to withdrawals of that funding, either because of generally stressed financial conditions or because of concerns with that particular bank. A bank might, therefore, choose to issue some long-term, floating-rate debt, thus obtaining some cushion of floating-rate funding that is much less subject to withdrawals. As it turns out, however, the market for such debt is extremely limited. A practical solution, therefore, is to create *synthetic* long-term, floating-rate debt, as illustrated in Figure 13.5. The bank sells long-term debt in the more traditional, fixed-rate market, and then receives fixed and pays floating in an IRS. The net result is long-term floating-rate debt, as desired.

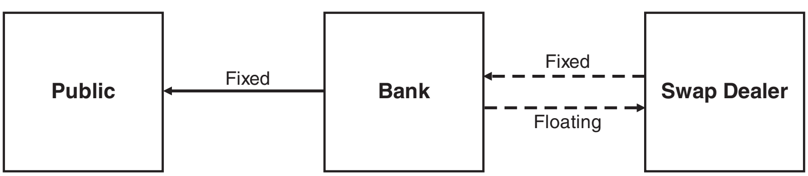

FIGURE 13.5 Synthetic Issue of Floating-Rate Debt.

13.4 COUNTERPARTY CREDIT RISK

Each counterparty to a swap contract bears the risk that the other will default on its obligations. A crucial mitigant of this risk is the *safe harbor* of swaps agreements from the bankruptcy code. In the event of a default, most creditors cannot immediately act to recover amounts owed to them. For example, a bank with an outstanding commercial loan that is secured by production machinery cannot, upon default, seize the collateral and sell it, and so recover the loan amount. Instead, the bank is subject to a bankruptcy *stay*. Under bankruptcy protection, the defaulting company is given time to reorganize, during which it may continue to use its machinery. At some point, with the permission of the bankruptcy court, the bank's loan will be paid off or restructured, possibly but not necessarily through the sale of the bank's collateral. Swap agreements, however, under their safe harbor, are not subject to the bankruptcy stay.[10] More specifically, if one counterparty defaults on its obligations under a swap contract, the other counterparty may terminate that contract and any others included in a master agreement with the defaulting counterparty; may net receivables and payables under all contracts included in the master agreement; and may liquidate any posted collateral under the agreement to cover swap closeout costs.

Without the safe harbor, then, the surviving counterparty would have to continue making the payments it owes, while no longer receiving payments from the defaulting counterparty, until a court settled the matter. With the safe harbor, however, the surviving counterparty may terminate contracts and stop making payments. Hence, the exposure to the surviving counterparty in the event of a default is the total NPV across all of the contracts with the defaulting counterparty. Put another way, any net positive NPV the surviving counterparty has against the defaulting counterparty is jeopardized by the default and termination of all trades. Furthermore, under the safe harbor, the surviving counterparty may recover potentially lost NPV by selling any collateral posted by the defaulting counterparty and, to the extent those proceeds are insufficient, may pursue an unsecured claim against the defaulting party through the courts. By the way, after a closeout, the surviving party must pay to settle any net negative NPV it has against the defaulting party. The policy justification behind the derivatives safe harbor is that financial institutions, which are typically highly leveraged, could suffer great losses through a bankruptcy process in which they did not know their risk, because they did not know which of their contracts would ultimately be honored

[10]Other financial transactions with a safe harbor include agreements to purchase or sell securities; various forward contracts for physical delivery; and repurchase agreements (repo) on government-guaranteed securities.

and which would not.[11] Since the financial crisis of 2007–2009, the safe harbor has been narrowed somewhat to give a governmental authority time to liquidate a failing, systemically important financial institution before its derivatives contracts are terminated, but that process is not explored further in this chapter.

One way dealers have managed the counterparty credit risk from derivatives customers is by collecting a fee, or insurance premium, sometimes called a *credit valuation adjustment (CVA)* charge, that incorporates the likelihood of default and the potential exposure of the position. These fees, often imposed on the customer as a higher rate when paying fixed or as a lower rate when receiving fixed, constitute a reserve that can make up for losses from the few defaults that actually occur. This means of managing credit risk is particularly suitable for a diversified group of creditworthy clients that have neither ready sources of cash nor the operational infrastructure to post and monitor collateral.

The predominant way dealers have managed counterparty credit risk, however, particularly when trading with professional investment firms, like asset managers and hedge funds, and with other dealers, is by taking collateral through *variation margin (VM)* and *initial margin (IM)*.[12] Calls for VM ensure that the counterparty with a positive NPV always holds sufficient collateral to cover a loss of that NPV, which, in this context, is sometimes called the *current exposure* of the swap. As an example, consider an IRS in which Counterparty A agrees to receive fixed from Counterparty B. At initiation, as discussed earlier, the NPV of a swap is zero. Now say that market interest rates fall such that Counterparty A has a positive NPV of $1 million against Counterparty B. In that case, Counterparty B must post $1 million of collateral to Counterparty A. Subsequently, however, rates rise dramatically, such that the NPV of the swap is positive $2 million to Counterparty B. Counterparty A must then return the $1 million that Counterparty B had posted and send an additional $2 million, so that now Counterparty B holds total net collateral from Counterparty A equal to its positive NPV or current exposure. VM calls are typically made daily, though they can be more frequent in times of heightened market volatility.

The VM arrangement described in the previous paragraph is now called VM *collateralized-to-market (CTM)*, because of a recent change in margin

[11] Derivatives markets in other countries have struggled without legal clarity on these matters. In fact, both India and China have recently taken steps to establish closeout netting in their systems. See Davis, C. (2021), "ISDA Poised to Issue India Netting Opinion," *Risk.net*, April 27, and ISDA (2021), "A Netting Milestone in China," ISDA, May 4.

[12] This discussion is similar to the discussion of collateral against repo transactions in Chapter 10.

arrangements for cleared IRS. For positions against a clearinghouse, IRS VM is now *settled-to-market (STM)*, which means that VM flows are not collateral postings, but irrevocable cash settlements of daily profit or loss on the position, just like the daily settlement payments of futures contracts described in Chapter 11. Recasting the example in the previous paragraph under VM STM, after rates fall, Counterparty B pays $1 million outright to Counterparty A. After rates subsequently rise, Counterparty A pays $3 million outright to Counterparty B. At that point, the NPV of the swap to Counterparty B is positive $2 million, but Counterparty B has already collected a net of $2 million from the VM STM payments.[13]

To complete the explanation of how VM protects the counterparty with positive NPV, continue the example by assuming that Counterparty A defaults right after its last VM payment. Under the safe harbor, Counterparty B tears up the swap with Counterparty A and, therefore, no longer receives or makes payments under that contract. Counterparty B needs to replace the contract, however, to restore the economics of the position before the default. Under VM CTM, Counterparty B relies again on the safe harbor to seize the net $2 million of collateral collected from Counterparty A through the VM arrangement, and pays that $2 million to some Counterparty C to enter into a swap at the same terms as the swap just canceled. By the definition of NPV – ignoring transactions costs, which are discussed further presently – Counterparty C is willing to receive fixed at a below-market rate in a swap with a negative NPV of $2 million in exchange for a payment of $2 million. Lastly, Counterparty C immediately sends that $2 million back to Counterparty B as collateral against its newly acquired negative NPV. To emphasize that it is fair for Counterparty C to enter into this swap, note that, while Counterparty C is receiving a below-market rate of interest, if rates remain constant, the NPV of the swap to Counterparty C will gradually increase from −$2 million to 0, and Counterparty C will get back $2 million of collateral that it never paid for. Put another way, by entering into this swap, Counterparty C essentially takes a position in a bond with a below-market coupon at a price of 98. The coupon is below market, but the price will increase to 100 at maturity.

Under VM STM, Counterparty B collected a total of $2 million over time as the NPV of the swap rose to $2 million, but holds no collateral at the time of default. Some Counterparty C again steps into the replacement swap, which receives a below-market rate and, consequently, has a negative NPV. Under VM STM, however, Counterparty C is responsible only for making payments arising from subsequent changes in NPV. Hence, there is no exchange of cash or collateral when Counterparty C enters into the swap

[13]The change from CTM to STM allowed the interest rate risk exposure of IRS of any maturity to be considered a one-day exposure for capital purposes.

under VM STM. Again, Counterparty C receives a below-market rate over time, but, on average, settlement payments are positive as the NPV rises from −$2 million to zero.

While VM is sufficient to protect a positive NPV immediately after each VM call, it does not protect changes in NPV between VM calls. Continuing with the example, say that rates rise again so that Counterparty B's NPV increases from $2 million to $3 million, and Counterparty A defaults before meeting its VM call. Counterparty B then has an exposure and possible loss of $1 million. To protect against value changes between VM calls, each counterparty holds a certain amount of IM posted by the other. The amount of IM is typically set so as to cover a large market move plus the transaction and liquidity costs of replacing the defaulting positions. For example, statistical analysis might indicate that, with 99% statistical confidence, the swap between Counterparty A and Counterparty B will move by less than $5 million before it can be replaced, and market expertise might indicate that replacing a swap of its notional amount incurs an additional $250,000 in transaction and market impact costs. The IM for the trade, therefore, might be set at $5,250,000. In this way, so long as the change in NPV between VM calls is less than posted IM, each counterparty has enough collateral to replace defaulted positions. Continuing with the example in a VM CTM arrangement, Counterparty B seizes $1 million of IM collateral, which, together with the VM of $2 million is enough to pay Counterparty C to take on the now negative $3 million NPV position. The remaining $4,250,000 of IM is returned to Counterparty A, the defaulting counterparty. In a VM STM arrangement, Counterparty B seizes $1 million of IM collateral, bringing the total collected to $3 million, which is now the NPV of the swap to Counterparty B. The remaining IM is returned to Counterparty A. Some Counterparty C steps into the swap, facing a negative NPV, but with obligations to make only future VM STM payments.

In general, collateral is posted in cash, in which case it earns some rate of interest, or in safe securities, like government bonds, in which case the counterparty posting the collateral keeps the interest. Securities posted as collateral might be accepted only at a haircut, that is, at less than their market value, so as to reflect the price risk should they have to be liquidated in the event of a default. At a haircut of 3%, for example, $100 of securities count only as $97 against collateral requirements. For cleared trades, the amount of IM to be posted is set and computed by the clearinghouse, and, for non-cleared trades, either using internal firm models or the industry *standard initial margin model (SIMM)*. An important quantity in the determination of IM is the *margin period of risk (MPOR)*, which is the assumed time interval over which a defaulted swap can be hedged or replaced. After a default, it is usual to hedge the replacement of the defaulted swap first and then – if desired – replace the actual swap with a willing counterparty and unwind the hedge. For example, say that Counterparty A defaults on its obligation to pay Counterparty B fixed at 2.34% on a swap with a remaining maturity of 12.3 years. Because it is likely to take some time to replace that particular swap, Counterparty B should first hedge the exposure it lost on account of the default by receiving fixed on a DV01-neutral amount of a 10-year swap at the prevailing market rate. Then, if desired, Counterparty B can find a counterparty to replace the original swap – along the lines previously discussed – and then, when the replacement swap is in place, unwind the hedge. In any case, the longer the MPOR, the greater the assumed possible changes in NPV before hedging or replacement, and the greater the required IM to ensure that sufficient funds are available for hedging and replacing the defaulted swap.

Before the financial crisis of 2007–2009, nearly all IRS were traded *over-the-counter (OTC)* and managed *bilaterally*. OTC trading means that the parties arrange transactions on their own, without a third-party platform or exchange. And in bilateral trades, each pair of counterparties sets margin rules, exchanges collateral, and bears the risk of each other's default. While IRS played no appreciable role in the financial crisis,[14] the Dodd-Frank Act required that all relatively liquid swaps be traded on a *swaps execution facility* and be cleared. The clearinghouse then sets margin rules and manages collateral, as is described in greater detail in the next section. Dodd-Frank further required that margin be exchanged by counterparties to any swaps that are not cleared, and that all swaps transactions and positions be reported to regulators through *swap data repositories*. Dodd-Frank and its implementing regulations exempt nonfinancial, *commercial end users*, who are using swaps to hedge, from clearing and margin requirements. This means, for example, that a nonfinancial corporation, which has neither the ability to fund margin calls nor the operational ability to manage the exchange of margin, can use a bilaterally arranged swap to hedge a debt issue, so long as it can find a willing dealer. As an aside, the business models of certain financial entities, like insurance companies and pension funds, are very suited to using IRS, but not to the liquidity demands of margin requirements.[15] These entities, however, unless they are very small, are not exempt from clearing and margin requirements.

[14]See Tuckman, B. (2015), "In Defense of Derivatives: From Beer to the Financial Crisis," *Cato Institute Policy Analysis*, Number 781, September 29.

[15]Consider a pension fund that receives fixed in IRS to hedge its long-term liabilities. With this hedge, the pension fund might very well be a safe swap counterparty, even without making VM payments. In fact, VM might increase the risk of the fund: suddenly rising rates would not change the net value of the fund – because it is hedged – but would result in VM calls that could require the pension to liquidate securities in a stressed environment. The fund could, of course, hold larger cash balances in readiness for such an eventuality, but those balances would result in a drag on returns.

The clearing requirement of Dodd-Frank resulted in the clearing of the vast majority of IRS, which, until recently, were dominated by LIBOR swaps. At the time of this writing, SOFR swaps are relatively new and have not yet been deemed liquid enough for the clearing requirement to apply. Nevertheless, over 70% of SOFR swaps are being cleared anyway, and the clearing requirement is likely to be applied to them in 2022 or soon thereafter.[16]

13.5 CLEARING AND CENTRAL COUNTERPARTIES

Figure 13.6 illustrates the difference between bilateral and cleared swaps. The top section of the figure shows a bilateral swap, in which Counterparty A pays fixed to and receives floating from Counterparty B. All aspects of the trade, from its execution to the ongoing exchange of cash flows and margin, are arranged between the two counterparties, and each counterparty bears all of the risk should the other default. While the diagram looks simple when there are only two counterparties, managing a large book of swaps trades with many counterparties is very complex. Every day, each counterparty has to send VM to or collect VM from each of its counterparties, in addition to sending or collecting any contractual interest payments due that day. Furthermore, each counterparty must track its exposure to each of its counterparties and perform ongoing due diligence as to their creditworthiness.

The middle section of Figure 13.6 illustrates a cleared swap between two firms that are members of the clearinghouse. Members 1 and 2 first execute a swap with each other, which is not shown in the figure, in which Member 1 pays fixed to Member 2. That swap is then *given up* for clearing, which means that it is canceled and replaced with the two swaps shown in the figure: one in which Member 1 pays fixed to the clearinghouse, acting in its capacity as a *central counterparty (CCP)*, and one in which the CCP pays fixed to Member 2. Note that the CCP takes no cash flow risk or market risk: the amounts received from Member 1 are simply passed to Member 2, and the amounts received from Member 2 are simply passed to Member 1. In any case, each member now legally faces the CCP and posts collateral to the CCP. If either member defaults, the CCP manages the default and suffers losses. However, as discussed further presently, large losses might have to be covered by the broader membership. This backstopping of losses by the membership is represented in the figure as the light gray circles surrounding the swap cash flows. On the other hand, if the CCP defaults, Members 1 and 2 have recourse only to the CCP, not to each other.

[16]St. Clair, B. (2021), "No Mandate, No Problem: SOFR Swaps Embrace Clearing," *Risk.net*, July 21.

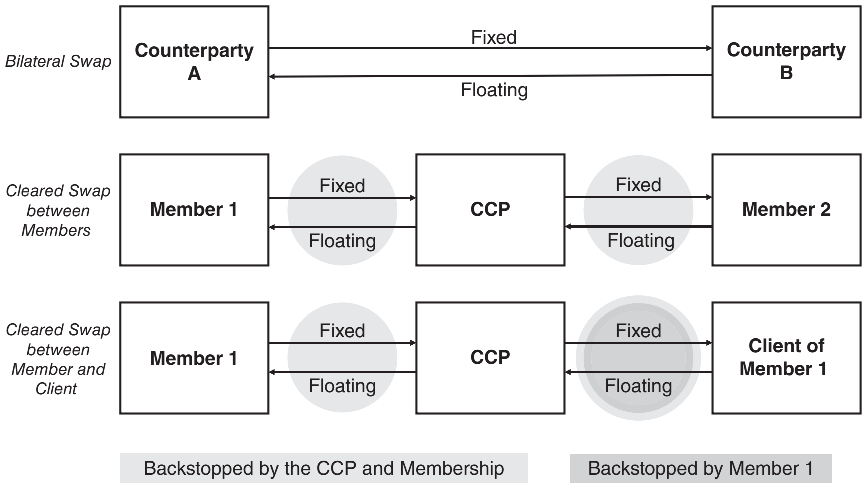

FIGURE 13.6 Bilateral versus Cleared Swaps.

The bottom section of Figure 13.6 illustrates a cleared swap between a member of the clearinghouse and a client of that member. The original swap between the two (not shown), as before, is transformed into two swaps, each facing the clearinghouse. And, as before, a default by the member is backstopped by the CCP and the broader membership. However, a default by the client, who is sponsored by Member 1 to face the CCP, is backstopped first by Member 1. This backstopping is represented in the figure by the dark gray circle around the cash flows of this swap. But if Member 1 defaults as well, then, as before, losses are backstopped by both the CCP and the broader membership. This fallback is represented in the figure by the light gray circle surrounding the darker one. While this section of the figure illustrates a trade between Member 1 and its client, the logic of this paragraph can be extended to other trading permutations: member trades are backstopped by the CCP and the broader membership, while client trades are backstopped first by the sponsoring member and then by the CCP and the broader membership.

Relative to bilateral trading of IRS, clearing greatly simplifies operations. Regardless of the number of positions and the history of who traded with whom, at the end of each day, each counterparty to the CCP makes only one net payment to or receives one net payment from the CCP. Clearing also greatly simplifies counterparty risk management in that each counterparty need be satisfied only with the creditworthiness of the CCP, although that creditworthiness – while extremely high – can be challenging to evaluate precisely. While clearing confers significant and extensive advantages, these come at some opportunity costs. Clearing nets positions within the same product class, but sacrifices netting across product classes. For example, a dealer who has IRS, repo, and credit default swap positions with one particular client might very well prefer to net the cash flows across those positions in-house instead of holding the IRS position against one CCP, the repo directly against the client, and the credit default swaps against a different CCP. The legal requirement to clear, however, outlaws this in-house netting option. Clearing also outsources margin methodologies and default management to the CCP, which, while advantageous to smaller counterparties, might not be optimal for those with broader operations.

Risk management at a CCP subsumes many functions. First, the CCP must set criteria for admitting members and monitoring their creditworthiness over time. Second, the CCP must set margin requirements. As discussed already, the amount of VM calls is conceptually straightforward, as they reflect changes in NPV. The appropriate amounts of IM, however, depend on complex analyses of market volatility and liquidity conditions. Third, the CCP must establish procedures for default management and execute them well as defaults occur. When a client of a member defaults, the sponsoring member is responsible, and when a member defaults, the CCP is responsible. More specifically, the CCP must make VM STM payments to the non-defaulting sides of the swap and replace the defaulting swaps along the lines described earlier. Note that, while a dealer may decide to replace lost exposure with similar exposure, the business model of the CCP is not to take any market risk at all. The swap lost must ultimately be replaced with a swap of exactly the same terms. In any case, the CCP's first source of funds is the IM posted by the defaulting member, along the aforementioned lines. If that proves insufficient, however, the CCP gathers the necessary funds according to a prespecified *default waterfall*.

Figure 13.7 illustrates a waterfall but is not drawn to the scale of the underlying resources. The first resource, just discussed, is the IM posted by the defaulting member. In addition to posting IM, however, members have to contribute to a *default fund* or *guarantee fund*, in proportion to the size of their positions against the CCP. One rule of thumb is that the total default fund should be large enough to withstand the simultaneous default of the two members with the largest positions. In any case, if the defaulter's IM is not sufficient to cover the CCP's losses, the defaulter's default fund contribution is tapped.[17] If that is insufficient as well, capital contributed by the CCP is tapped. Relative to the size of the total of all members default fund

[17]In managing defaults on cleared IRS, the margin of non-defaulting customers of members is not legally at risk. Their margin is *legally separated operationally commingled (LSOC)*. This means that, while their margin might be mixed with that of other customers for operational convenience, if the member defaults with an outstanding obligation to the CCP, the margin of the non-defaulting customers of that member are not legally available to the CCP. This treatment differs from that of customers in the futures market, where, should their *futures commission merchant*

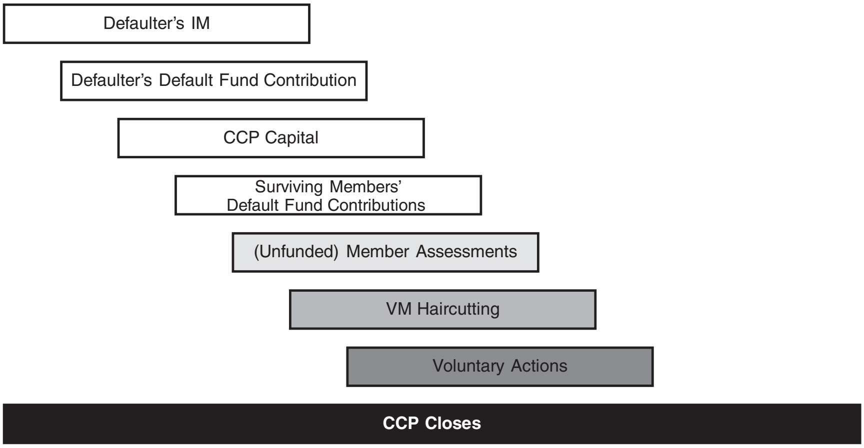

FIGURE 13.7 Example of a CCP Default Waterfall. Steps Are Not Drawn to Scale.

contributions, CCP capital is typically quite small. This buffer is sometimes referred to as the CCP *skin-in-the-game*, and its size is the subject of much debate. To oversimplify that debate, CCPs argue that, because almost all of the risk comes from member positions, almost all of the waterfall protections should come from members. Members, on the other hand, argue that, because CCPs manage risk with the profits of their shareholders as an important consideration, they should contribute a significant buffer as well. In any case, if CCP capital is exhausted with losses still needing to be covered, the default fund from the surviving members is tapped. The waterfall to this point explains the sense in which the membership as a whole backstops member contracts. It is also the sense in which CCPs *mutualize* losses. Loose descriptions of clearing say that CCPs "eliminate" counterparty credit risk, but that is not accurate. When a swap is cleared, its counterparty credit risk moves from the original counterparty to the CCP and its membership as a whole.

With VM covering daily market moves, IM covering all but the largest of market moves between VM calls, and the default fund sized to withstand the default of the two members with the largest positions – who are almost certainly large and heavily regulated financial institutions themselves – only a record-breaking financial tsunami could generate losses that require stepping further down the waterfall. The next sources of funds, however, are as follows. Members agree to comply with certain assessments or call for

---

(FCM) default with an obligation to the futures clearinghouse, bankruptcy proceedings could make non-defaulting customer margin available to the clearinghouse. This possibility is known as *fellow customer risk*.

funds should the waterfall get to this point. These assessments are called *unfunded*, because the CCP does not hold them in reserve, as it does IM and default fund contributions. While legally binding, there is always some concern that, in a crisis, assessments might not be honored promptly and in full. If assessments are insufficient to cover losses, then some fraction of VM owed to market participants would not be paid, in a process known as *VM haircutting*. After that, the sources of last recourse include voluntary contributions and other actions taken by members. If that proves insufficient, CCP ceases operations.

This section concludes with a brief mention of three topical public policy issues. First, the shift to clearing the vast majority of IRS may or may not have significantly reduced systemic risk relative to various alternatives, but it has certainly concentrated that risk. Nearly all GBP-denominated IRS are cleared at the London Clearing House (LCH); the vast majority of USD-denominated IRS are cleared at LCH as well, with the Chicago Mercantile Exchange (CME) a distant second; the vast majority of EUR-denominated swaps are cleared at LCH, with the exception of the smaller but growing segment of €STR swaps, in which Eurex has a growing market share; and clearing of JPY-denominated swaps is shared between LCH and the *Japanese Securities Clearing Corporation (JSCC)*.[18] From a systemic risk perspective, legislators and regulators seem to have followed the advice of Andrew Carnegie: "put all your eggs in one basket and then watch that basket."

A second policy issue pertains to CCP margin. Margin, along with other risk management practices of a CCP, are designed to make it extremely unlikely that any counterparty loses money from counterparty risk. Part of this outcome, however, is due to the ability of CCPs to set IM and also to raise required IM when market volatility or financial stress increases. In this sense then, reducing counterparty risk increases liquidity risk. Many are concerned with how increasing IM might exacerbate financial stress, known as *margin procyclicality*, but there are no easy solutions. Raising margin in a crisis is an important risk management tool at a CCP but might make it challenging for members and other market participants both to meet their increasing obligations to the CCP and to continue meeting their non-derivatives obligations.

The third policy issue relates to CCP governance. Before clearing was required, market participants could consider clearing as one of several

---

[18]Khwaja, A. (2021), "2020 CCP Volumes and Market Share in IRD," *clarusft.com*, January 13. As an aside, clearing of credit default swaps is also concentrated, with nearly all of USD volume at the Intercontinental Exchange (ICE) and EUR volume split with very roughly 80% at ICE and 20% at LCH. Khwaja, A. (2022), "2021 CCP Volumes and Shares in CRD," *clarusft.com*, February 1.

choices and could decide to clear or not to clear. With required clearing and the concentration of clearing in very few CCPs, market participants, including members, have much less leverage with respect to CCP risk management and other practices. The resolution of this issue is also far from straightforward.

13.6 BASIS SWAPS

Both OIS and fixed-for-floating swaps pay interest at a fixed rate on one leg and at a floating rate on the other leg. A *basis swap*, by contrast pays interest at one floating rate on one leg and interest on a different floating rate on another leg. Before the transition away from LIBOR, basis swap volumes were particularly large in exchanging LIBOR of one term against LIBOR of another term, and swaps of effective fed funds (see Chapter 12) against LIBOR of various terms were common as well. In US dollar markets, swaps of SOFR against LIBOR are trading actively in the transition period but will fade with the disappearance of LIBOR. *Cross-currency basis swaps* have been and continue to be extremely popular. These swaps exchange interest at a short-term rate in one currency for interest at a short-term rate in another currency, for example, SOFR versus €STR. This section focuses on basis swaps of €STR versus three-month Euribor to explain concepts, both because Euribor is still an active rate and because foreign exchange rates are outside the scope of this edition of the book.

Consider a bank that funds itself at €STR flat (i.e., without a spread) and lends money to customers at three-month Euribor plus 1%. This bank has the *basis risk* that €STR, compounded over a relevant time period, rises relative to three-month Euribor, compounded over that same period. In general, by the way, the term "basis risk" refers to the risk that two rates or prices, which usually change together in some fixed or predictable relationship, diverge in an unusual and unfavorable way. In any case, Figure 13.8 shows how the bank can hedge its basis risk with an €STR versus three-month Euribor basis swap. As of February 2022, the two-year €STR versus three-month Euribor *basis swap spread* was 13.8 basis points, which means that the bank, as shown in the figure, can receive €STR plus 13.8 basis points and pay three-month Euribor on some notional amount for two years. More specifically, at the end of every quarter during those two years, the bank receives interest at daily compounded €STR plus 13.8 basis points over that quarter (along the lines of Figure 13.1), and the bank pays interest at three-month Euribor set at the start of the quarter (along the lines of Figure 13.2). Overall then, with its basis swap hedge, the bank in Figure 13.8 locks in a fixed rate of 1.138%, so long as its customers do not default. Put another way, the customers pay a credit and liquidity spread of 1.138% over the near-riskless

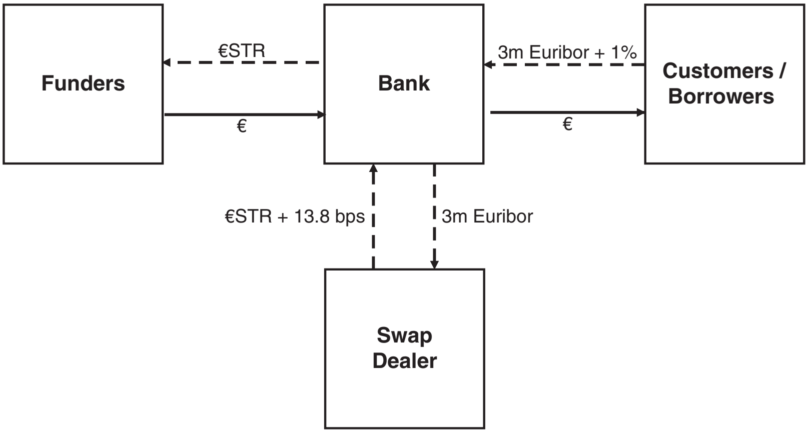

FIGURE 13.8 Hedging Basis Risk with an €STR versus Three-Month Euribor Basis Swap.

€STR: 1% as a spread over three-month Euribor and 13.8 basis points as a spread of three-month Euribor over €STR.

From the swap dealer's perspective, there is virtually no credit risk: the bank itself is probably a good credit and the basis swap is collateralized. Because Euribor is a riskier rate than €STR, however, compounded €STR over a quarter will normally be less than three-month Euribor. It makes sense, therefore, that paying a spread over €STR is fair against receiving three-month Euribor, although the market determination of the exact spread can certainly be the subject of further analysis.

The existence of basis swap spreads shows that not all swaps can be valued using the methodology for pricing swaps that is given in Chapter 2 and earlier in this chapter. That methodology argued that the floating leg (including the fictional notional amount) is worth par when the floating rate is considered to be the risk-free rate, that is, the rate at which funds can be moved without risk across time. And it is reasonable to classify both overnight SOFR and overnight €STR as risk-free rates.[19] The same is not true, however, for three-month Euribor. Or, to put it another way, it

[19]This sentence means that swaps in USD price the floating leg of a SOFR swap at par and that swaps in EUR price the floating leg of an €STR swap at par. However, in a *cross-currency basis swap* of SOFR *versus* €STR, only one of the floating legs can be worth par. Furthermore, there is a *cross-currency basis swap spread* added to one of the floating legs to make the swap fair. Conceptually, if the collateral posted against the swap earns SOFR, then SOFR is the risk-free rate for valuation purposes, and the SOFR floating leg is worth par.

cannot be simultaneously true that i) a floating leg paying €STR is worth par; ii) €STR plus 13.8 basis point is fair against three-month Euribor for two years; and iii) a floating leg paying three-month Euribor for two years is worth par.

The appropriate methodology for valuing a fixed-for-floating swap when the floating index is not a risk-free rate, like Euribor, is the following. First, discount the fixed cash flows (including the fictional notional amount) at the discount factors implied by swaps against the risk-free rate index, like €STR. These calculations are described in Chapter 2. Second, the value of the floating leg is equal to par plus the present value of payments of the basis swap spread, where discounting is again done using the risk-free rates. For example, the basis swap in Figure 13.8 shows that the value of receiving three-month Euribor equals the value of receiving €STR plus 13.8 basis points. But the value of receiving €STR (including the fictional notional amount) is par. Therefore, the value of receiving three-month Euribor quarterly for two years is par plus the present value of receiving 13.8 basis points, paid quarterly, for two years.

Appendix A13.1 illustrates the calculations outlined in the previous paragraph with representative market rates toward the end of February 2022, when the fixed rate on a two-year, fixed versus three-month Euribor swap was 0.078% and, as already mentioned, the two-year €STR versus three-month Euribor basis swap spread was 13.8 basis points. The values of both the fixed and floating sides are calculated to be 100.281. To summarize, when the floating index is a risk-free rate, both sides of a fixed-for-floating swap at initiation are worth par. When the floating index is not a risk-free rate, but trades with a positive basis swap spread against the risk-free rate, both sides of the swap at initiation are equal in value, of course, but each side is worth more than par. Note that traders may just accept that rates on newly initiated swaps are fair and, therefore, may not care about the value of each side separately. Nevertheless, a methodology is required to compute the NPVs of existing swaps in a trading book, which are not observable in the market.

The pricing methodology just described can take rates on €STR swaps and spreads on €STR versus three-month Euribor basis swaps as given to price fixed versus three-month Euribor swaps. Alternatively, the methodology can take rates on €STR swaps and on fixed versus three-month Euribor swaps as given to imply fair basis swap spreads. The latter approach, called *two-curve pricing*, is often preferred when basis swaps of all terms are insufficiently liquid. In fact, two-curve pricing typically does not explicitly calculate basis swap spreads at all. A brief description of this methodology is given in Appendix A13.2.

CHAPTER 14

Corporate Debt and Credit Default Swaps

Corporations borrow money to fund their operations, transactions (e.g., mergers and acquisitions), and changes to capital structure (e.g., refinance existing debt, stock repurchases). The loans and bonds used to raise these funds are subject to *credit risk*, because corporations may not make good on their promises to pay interest and repay principal. Lenders, in turn, require compensation for bearing credit risk in the form of higher returns. The cash flows of *credit default swaps* (CDS) also depend on the payment or non-payment of debt obligations but are not themselves obligations of corporate issuers. In other words, through CDS contracts, market participants trade corporate credit risks with each other. While this chapter focuses mostly on corporate debt, much of the discussion applies to sovereign and municipal debt as well, because the debt obligations of many governments have at times been perceived as subject to nontrivial probabilities of default.

Many market participants rely to some extent on *rating agencies* to measure the credit risk of borrowers and their loans and bonds. The long-term debt ratings classifications of the three major rating agencies in the United States are given in Table 14.1. More granular breakdowns of each of these broad ratings classifications are also available, and short-term debt has its own, separate scales.[1]

[1]Moody's Aa ratings, for example, are divided into Aa1, Aa2, and Aa3, and its B ratings into B1, B2, and B3. S&P and Fitch AA ratings are divided into AA+, AA, and AA-, and their B ratings into B+, B, and B-. Short-term debt ratings are P-1, P-2, P-3, and "Not Prime" by Moody's; A-1+, A-1, A-2, and A-3 by S&P; F1+, F1, F2, and F3 by Fitch; followed by B, C, and D for both S&P and Fitch.

TABLE 14.1 Long-Term Debt Ratings Classifications. The Ratings in Each Entry Are Listed in Decreasing Order of Creditworthiness.

Agency	Investment Grade	Speculative Grade/ High Yield	Default
Moody's	Aaa, Aa, A, Baa	Ba, B, Caa, Ca	C
S&P, Fitch	AAA, AA, A, BBB	BB, B, CCC, CC, C	D

14.1 CORPORATE BONDS AND LOANS

Large and highly creditworthy corporations in the United States tend to borrow money for fixed terms in public markets through *commercial paper* (CP), *medium-term notes* (MTNs), and corporate bonds. CP is typically a discount (i.e., zero coupon) instrument that is either unsecured, backed by a letter of credit from a bank, or secured by assets. CP is exempt from Securities and Exchange Commission (SEC) registration, along with its associated costs, so long as the paper matures in less than 270 days and its proceeds are used for short-term purposes (rather than, for example, building a factory). With its high credit quality and short-term maturity, CP is a particularly inexpensive and liquid way for the most creditworthy corporations to raise short-term funds. CP borrowing does, however, expose a corporation to the funding or liquidity risk of having to roll outstanding CP as it matures into new CP issues.

Corporations sell MTNs in public markets to raise money with customized payment terms. Historically, MTNs primarily filled the maturity gap between CP and longer-term bonds but are now better characterized as debt instruments customized to suit the needs of individual issuers and investors. MTNs first became popular in the United States in the early 1980s, with the introduction of SEC shelf registrations. These programs allow issuers to register once to sell notes opportunistically, over time, at terms that can be adjusted at the time of each sale.

For longer-term borrowing in public markets with relatively standard payment terms, corporations sell bonds, which, in the United States, have to be registered with the SEC. Corporate bonds are typically coupon-bearing, fixed-rate securities, but there is a much smaller market for *floating-rate notes* (FRNs). The interest rate on FRNs is typically set equal to a short-term reference rate plus a fixed spread, although a reference rate might be multiplied by a factor or *leverage*, and the spread might vary over time with the credit rating of the issuer. The short-term reference rate for FRNs had traditionally been the *London Interbank Offered Rate* (LIBOR) but is now transitioning to the *Secured Overnight Financing Rate* (SOFR) and other non-LIBOR alternatives (Chapters 12 and 13).

Smaller and less creditworthy corporations, which cannot typically raise funds in public markets, tend to borrow money through *private placements* of bonds and through bank loans. In a privately placed bond issue, the borrower tailors terms to satisfy its own requirements and those of a relatively small group of lenders. Insurance companies are the most significant investors in this market, though other asset managers participate as well. Privately placed bonds are exempt from SEC registration precisely because they may be sold only to investors deemed "sophisticated."

Bank loans tend to be floating-rate instruments, although borrowers sometimes convert their debt into a fixed rate by paying fixed to and receiving floating from a bank in an accompanying interest rate swap, which the bank, in turn, hedges with a dealer (Chapter 13). From the bank perspective, floating-rate loans have the advantage of matching floating-rate liabilities, which are mostly deposits, but also include wholesale funding like commercial paper. Also, floating-rate loans are the easiest to sell in the secondary market for bank loans.

Traditionally, banks made loans and held them to term. And if a single borrower needed a loan that was too large for one bank, either because funds were not readily available or, more likely, because the resulting credit risk would be too significant for one bank, several banks would form a group or *syndicate* in which each would make and hold a smaller loan to that large borrower. The last few decades, however, have seen phenomenal growth in the secondary market for loans, which allows banks to sell loans they have made to institutional investors. In this way, banks can earn fees on making, servicing, and monitoring loans without having to warehouse all of the associated credit risk. At present, in fact, the overwhelming majority of relatively low-quality or *leveraged* loans are held not by banks, but by institutional investors. While some leveraged loans are bought from banks directly by insurance companies, mutual funds, and hedge funds, much of the growth of the secondary market for bank loans has been through the indirect sale of loans to institutional investors through *collateralized loan obligations* (CLOs).

Collateralized Loan Obligations

A CLO is a vehicle that purchases a portfolio of leveraged loans financed by the sale of several debt classes or *tranches* and equity or subordinated notes. Table 14.2 lists the debt tranches of a particular CLO issued in May 2019. Essentially, interest and principal payments from the underlying portfolio of leveraged loans flow to the tranches from the top down, while any credit losses are allocated from the bottom up.[2] Senior Secured Floating-Rate Note

---

[2]The principal of maturing loans in this CLO is reinvested through April 2024. From then on, however, maturing loan principal is passed to the tranche holders. CLO

TABLE 14.2 Tranches of Apidos CLO XXXI, May 2019. All Tranches Mature in April 2031.

Class	Description	Amount ($millions)	Rating S&P/Moody's	Spread over LIBOR (bps)
X	Sr. Secured FRNs	4.725	AAA/–	65
A-1	Sr. Secured FRNs	370.150	AAA/Aaa	133
A-2	Sr. Secured FRNs	39.350	–/Aaa	165
B	Sr. Secured FRNs	64.600	AA/–	190
C	Mezzanine Secured Deferrable FRNs	40.300	A/–	255
D	Mezzanine Secured Deferrable FRNs	36.200	BBB–/–	365
E	Mezzanine Secured Deferrable FRNs	27.400	–/Ba3	675
–	Subordinated Notes	51.400	–	–
	Total	634.125		

Classes X, A-1, A-2, and B are paid interest and principal from the underlying loans, while Mezzanine Secured Deferrable classes C, D, and E are paid to the extent that additional cash flows from the underlying loans are available. Any credit losses from the underlying loans are first applied to the equity or Subordinated Notes, then to the Mezzanine tranches, and only then, if necessary, to the Senior Secured tranches. Furthermore, CLO debt tranches are protected by constraints imposed on the creditworthiness of the underlying loans and on the concentration of loans to a particular borrower or industry, along with various ongoing tests as to the sufficiency of interest and collateral. In short, this CLO, from an underlying portfolio of $634,125,000 leveraged loans, created $414,225,000 of X, A-1, and A-2 tranches with AAA or Aaa ratings; an additional $64,600,000 B tranches with an AA rating; and so forth.

Compensation for buying lower-rated tranches comes in the form of earning a higher spread. As shown in the last column if Table 14.2, spreads range from 65 basis points for the AAA-rated X Class to 675 basis points for the Ba3-rated Class E tranche.[3] The return on equity, of course, is determined by whatever is left over after paying all of the more senior claims.

The different credit risk profiles of the different tranches in a CLO tend to attract different groups of investors. Banks tend to hold or purchase the AAA/Aaa and AA/Aa tranches; insurance companies and pension funds invest in a range of tranches; and hedge funds and alternative asset managers

---

managers here, therefore, reinvest loan principal so that loans in the portfolio as of April 2024 mature at or somewhat before the maturity of the tranches in April 2031.

[3]In the case of this CLO, the X Class is particularly short term and safe, because its principal begins to amortize soon after issuance.

buy the lower-rate tranches. Outside investors and the CLO originators hold the equity.

By way of summarizing trends over the last few decades, the growth of the secondary market for leveraged loans, abetted by the CLO market, has blurred the distinction between leveraged loans and high-yield bonds. Corporations have greater flexibility to borrow in one market or the other, and asset managers actively decide to invest in one market or the other, with all choices depending on individual preferences or requirements and on market conditions.

Seniority, Covenants, and Call Provisions

When selling a debt issue, a corporation enters into a contract with bondholders, called an *indenture*, which is enforced by a trustee. Aside from payment terms (e.g., interest, principal, and maturity), the indenture specifies the priority of the issue in the event of default. For example, one bond issue might be secured by a particular set of assets; a second might be unsecured, but "senior" to other issues; and a third might be "subordinated" to higher-ranking issues. In this example, should a corporation be reorganized or liquidated through bankruptcy according to "strict priority," proceeds from selling ring-fenced assets would be applied first to satisfy the claims of the secured bondholders. Any remaining proceeds, together with other assets of the corporation, would be applied next to satisfy the claims of the senior bondholders. Finally, whatever value remains would be used to satisfy the claims of the subordinated bondholders. In practice, because reorganizations involve negotiation and include equity holders, strict priority may not always precisely predict the settlement of debt claims.

Indentures also include *covenants* to protect the claims of bondholders. Examples include maintaining various financial ratios; restricting the amount of cash that can be paid to stockholders; requiring a corporation to repurchase a debt issue after a change of control; limiting the total amount of new debt incurred by the corporation; and preventing the sale of debt with higher seniority than that of an outstanding debt issue.

As a final point of discussion, indentures often include *call provisions* or *embedded* call options, which allow a corporation to repurchase bonds from bondholders at some fixed schedule of prices. A simple example would be a 20-year bond that the company can repurchase at par or face value anytime after 10 years. A more complex example would be a 4% 30-year bond that the company can repurchase after 10 years at a price of 102 (100 plus half the coupon); at a price of 101.90 after 11 years; at a price of 101.80 after 12 years.

The original purpose of call provisions was to enable corporations to extinguish a bond issue – without having to track down and purchase every outstanding bond – perhaps to remove covenants that prohibited what had

TABLE 14.3 Call Provision of the Hertz 6s of 01/15/2028. The Bond Was Issued in November 2019.

Dates	Call Price
Before 01/15/2023	Make-Whole Price @Yield of the Treasury 2.75s of 02/15/2028 +50bps
01/15/2023 to 01/14/2024	103.00
01/15/2024 to 01/14/2025	101.50
On or after 01/15/2025	100.00

become value-added activities, or perhaps to change the corporation's existing capital structure. As interest rates became more volatile in the early 1980s, these call provisions became valuable as interest rate options: the value of the right to purchase a bond at some fixed price increases as rates decline (bond prices increase) and also as rate volatility increases. At the same time, of course, corporations issuing debt with call provisions have to pay investors for that option through a higher coupon rate. Chapter 16 describes the pricing of these call provisions.

Since the mid-1990s, however, the most common call provision has become the *make-whole* call. For example, in a private placement in November 2019, Hertz sold the 6s of 01/15/2028, with the make-whole call provision described in Table 14.3. Before January 15, 2023, Hertz can repurchase the bonds from investors at a yield 50 basis points above the yield of the US Treasury 2.75s of 02/15/2028. From January 15, 2023, on, the call is at the schedule of fixed prices shown in the table. The idea behind the make-whole call is to give the issuer flexibility to manage its debt without its having to purchase an expensive interest rate option. Unlike the case of a call price that is fixed, when yields fall, the yield of the Treasury benchmark falls too, and the make-whole call price increases. Hence, the interest rate option value of the make-whole call is extremely limited. In fact, in theory, the make-whole call has option value only if the market spread of the Hertz bonds over Treasuries falls: in that scenario, the value of the Hertz bond increases without a corresponding fall in the yield of the Treasury bond and a corresponding increase of the make-whole price. In practice, however, the value of this spread option is mitigated by setting the make-whole spread out-of-the-money. For example, when Hertz issued the 6s of 01/15/2028 in November 2019, its spread over the Treasury benchmark was 423 basis points. This means that its bond spread would have to fall a very large 423 – 50 or 373 basis points before the make-whole provision was in-the-money as an option on spread. Overall, of course, the call provision of the Hertz 6s of 01/15/2028 derives option value from the schedule of fixed call prices from January 15, 2023.

14.2 DEFAULT RATES, RECOVERY RATES, AND CREDIT LOSSES

While investors in nearly risk-free government bonds can focus exclusively on interest rate risk, investors in corporate securities have to focus on credit risk as well, often expressed in terms of *default rates*, *recovery rates*, and *credit losses*. After analysis, an investor might estimate that, over a five-year horizon, a particular portfolio of bonds will experience a default rate of 10%, meaning that, for every $100 of face amount, $10 will default. Furthermore, the investor might estimate that the defaulting bonds in the portfolio will experience a recovery rate of 40% of face amount, or, equivalently, will suffer a loss equal to the remaining 60% of face amount. Putting these two estimates together, the investor expects, over a five-year horizon, that credit losses on the portfolio will be $10\% \times (1 - 40\%) = 6\%$.

Table 14.4 shows average historical values for these quantities, over the period 1983–2020, for senior unsecured bonds, by rating. The results are useful for appreciating the average magnitude of credit risk. For investment-grade senior unsecured bonds, the average five-year default rates and recovery rates are 0.9% and 44.5%, respectively, giving an average credit loss of 0.5%. For speculative-grade senior unsecured bonds, the average default rate is much higher, at 19.6%, though the average recovery rate is only marginally lower, at 38.3%. Combining these two averages gives a much higher credit loss of 12.2%. The standard assumption in the industry that recovery rates are about 40% is justified by historical data like those presented in Table 14.4.[4]

While historical averages are useful in thinking about credit risk, credit conditions can vary dramatically over time. For example, extending the sample of Table 14.4 to include the years after the Great Depression raises the five-year default rate on investment-grade debt from 0.9% to 1.4%. And

TABLE 14.4 Average Five-Year Default Rates, Senior Unsecured Bond Recovery Rates, and Credit Losses, 1983–2020. All Entries Are in Percent.

Rating	Default Rate	Recovery Rate	Credit Loss
Investment Grade	0.9	44.5	0.5
Speculative Grade	19.6	38.3	12.2
All	7.4	38.9	4.6

Source: Moody's Investors Service

[4]These data calculate recovery rates from bond prices just after default, as opposed to ultimate recovery rates at the conclusion of reorganizations or liquidations.

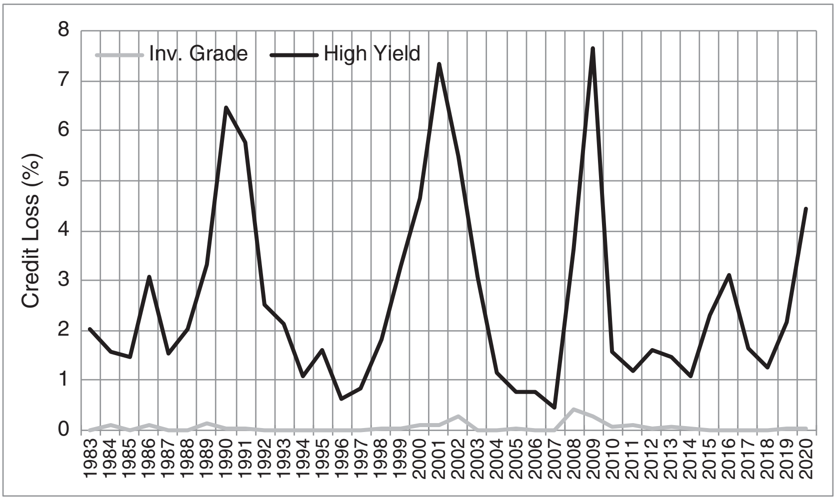

FIGURE 14.1 Credit Losses for Senior Unsecured Bonds, 1983–2020. *Source:* Moody's Investors Service.

within the sample period of the table, Figure 14.1 shows the variability of credit losses for investment-grade and high-yield bonds. High-yield losses were particularly high in the late 1980s, following the rapid growth in that market; in 2000–2002, which included the "dot-com" crash and the failures of Enron and WorldCom; in the financial crisis of 2007–2009; in 2016, from stresses caused by low prices in energy markets; and, most recently, during the pandemic and economic shutdowns of 2020.

Table 14.4 and Figure 14.1 report credit losses for senior unsecured bonds. Investment outcomes vary significantly, however, across loans and bonds of different seniority. Over the same sample period as in the table and figure, Moody's reports that bond recovery rates were 22% for junior subordinated; 38% for senior unsecured; and 54% for first lien (i.e., the highest priority secured claim). For bank loans, average recovery rates were 46% for senior unsecured; 65% for first lien; and 32% for second lien.

A recent example of the impact of seniority, illustrated in Table 14.5, is the price behavior of Hertz bonds through its bankruptcy filing on May 22, 2020. Before the pandemic, as of February 21, 2020, when default was a remote contingency, the prices of Hertz' secured and unsecured bonds were above par, with the higher price of the unsecured reflecting its longer maturity at an above-market coupon. By May 14, 2020, after the pandemic and shutdowns devastated the rental car business, the prices of Hertz bonds plummeted, with the price of the secured bond now significantly higher: with default imminent, the seniority of the secured bonds was much more important than the seemingly distant prospect of the unsecured bond's cash

TABLE 14.5 Hertz Corporation, Selected Bond Prices on Three Dates in 2020.

Priority	Coupon	Maturity	Feb 21	May 14	Jun 15
Sr. Secured 2nd Lien	7.625	06/01/2022	103.10	19.97	77.00
Sr. Unsecured	6.000	01/15/2028	104.29	11.25	43.50

flows through 2028. Soon after, however, in the wake of the economy's rapid recovery, Hertz's bond prices recovered dramatically as well. There was still much uncertainty as to the bonds' creditworthiness, however. The June 15, 2020, prices in the table show that the secured bonds still sold at a significant premium to the unsecured. By way of epilogue, Hertz emerged from bankruptcy at the end of June 2021, and its bonds recovered their full principal value.

Potential credit losses are directly relevant for investors expecting to hold bonds to maturity. Investors with shorter horizons, however, are also concerned about credit deterioration, which causes bond prices to fall before a realized default and credit loss. One manifestation of credit changes are rating *transitions*, in which a rating agency upgrades or downgrades a rating. Table 14.6 gives average historical one-year transition rates, or, more specifically, the rates at which bonds starting at a particular rating, over the following year, are upgraded, experience no change, are downgraded (including defaults), or become unrated. Becoming unrated may be a negative credit event but could also indicate a credit-neutral event like an acquisition. In any case, the table shows that downgrades over a one-year horizon are not uncommon, averaging from about 4% to 10% for bonds rated B or above, and over 28% for CCC/C-rated bonds. Upgrades occur as well, but with less frequency. Though not shown in the table, upgrades and downgrades vary with the business cycle, like the credit losses in Figure 14.1.

TABLE 14.6 Average One-Year Transition Rates, 1981–2020. All Entries Are in Percent.

Rating	Upgrade	No Change	Downgrade	No Rating
AAA	0.0	87.1	9.8	3.1
AA	0.5	87.2	8.4	3.9
A	1.6	88.6	5.4	4.4
BBB	3.3	86.5	4.3	5.9
BB	4.7	77.8	8.0	9.5
B	4.8	74.6	8.3	12.3
CCC/C	13.3	43.1	28.3	15.3

Source: S&P Global Ratings; and Author Calculations

14.3 CREDIT SPREADS

Credit spreads are the differences between the relatively high rates earned on fixed income instruments that are subject to credit risk and the relatively low rates on instruments with little or no credit risk. The simplest measure of credit spread is the *yield spread*, which is the difference between the yield on a bond and the rate or yield on a similar maturity interest rate swap or highly creditworthy government bond. Yield spreads, however, suffer from a number of drawbacks. First, a sufficiently liquid government bond or swap with a similar maturity might not exist. Second, yields reflect not only credit risk, but also the structure of a bond's cash flows. (See the discussion of the "coupon effect" in Chapter 3.) Third, yields reflect the value of embedded options, like the fixed-price call provisions discussed earlier, which have nothing to do with credit risk.

A better measure of credit spread is referred to in this chapter as the *bond spread*. This term includes the spread defined in Chapter 3 and the more general *option-adjusted spread (OAS)* defined in Chapter 7. In the credit context, a bond spread is computed by assuming no default and finding the spread over a benchmark curve that prices a bond as it is priced in the market. Because the market price incorporates the risk of default while this pricing methodology does not, the bond spread is a metric of credit risk. Unlike yield spreads, bond spreads properly account for maturity and the structure of cash flows. This approach, therefore, is suitable for bonds without embedded options. For bonds with embedded options, OAS is more appropriate: by design, its computation of price accounts for the value of embedded options and, therefore, any remaining spread to the benchmark curve can be reasonably attributed to credit risk. Furthermore, as explained in Chapters 3 and 7, the bond spread and OAS can be interpreted as the extra spread earned by a bond if interest rates are unchanged or hedged against; if the spread is unchanged; and if the bond does not default.

The measures of credit spread in this section are illustrated with the Genworth 4.90s of 08/15/2023, a speculative-grade bond issued by an insurance company. Table 14.7 gives various measures of credit spread for this bond, as of August 15, 2021, when the 4.90s of 08/15/2023 have exactly two years to maturity. Given the market yield of 5.596% and the two-year LIBOR swap rate of 0.288%, the yield spread is the difference between those two rates, or 530.8 basis points. The bond spread, given the bond's market price of 98.70 and the term structure of forward swap rates as of the pricing date, is 531.1 basis points.[5] Then, in the sense of the previous paragraph, the Genworth bond earns an annual rate of LIBOR plus 531.1 basis points.

[5]For readers wanting to follow these calculations, the first four six-month forward rates from the swap curve were taken to be 0.154%, 0.154%, 0.4225%, and 0.4225%.

TABLE 14.7 Selected Credit Spreads for the Genworth 4.90s of 08/15/2023, as of August 15, 2021. The Price of the Bond is 98.70; Its Yield is 5.596%; and the Par Swap Rate Is 0.288%. Spreads Are in Basis Points.

Spread Type	Spread
Yield Spread	530.8
Bond Spread	531.1
Par-Par Asset Swap Spread	526.4
Market Value Asset Swap Spread	533.3

Other popular measures of spread are *asset swap spreads*. The point of an *asset swap* is to transform a fixed-rate, coupon bond into an asset that earns a spread over a short-term rate, like LIBOR. In the context of this chapter, asset swaps enable investors to earn credit spreads without having to bear the interest rate risk of long-term, fixed-rate bonds.

Figure 14.2 illustrates one type of asset swap, the *par-par asset swap*, or, more simply, the *par asset swap*. Thin lines indicate cash flows at the initiation of the asset swap; heavy lines indicate intermediate cash flows; and dashed lines indicate cash flows at the maturity of the swap. At initiation, the asset swapper buys the bond for P per 100 face amount, earning a periodic coupon payment of c. The purchase price is financed with 100 from the repo desk and $P - 100$ from the swap desk.[6] Finally, through an interest rate swap to the maturity of the bond, in addition to the up-front payment just mentioned, the asset swapper periodically pays c in exchange for receiving LIBOR plus the spread, s^{par}, on 100. Note that this trade requires no cash at initiation; earns LIBOR plus s^{par} minus the repo rate over the life of the bond, so long as the bond does not default; and requires no cash at maturity, again, so long as the bond does not default. Hence, as desired, the investor has converted the fixed coupon payments of a bond into floating payments of LIBOR plus s^{par}. Of course, there are losses if the bond defaults: in that case, the asset swapper has to make coupon payments to the swap desk and, at maturity, pay 100 to the repo desk, even though these payments will not be fully realized from the defaulted bond.

The par asset swap spread, s^{par}, can be determined by the condition that the interest rate swap with the swap desk is fair, that is, that the initial payment plus the present value of the floating leg equals the present value of the fixed leg, where discounting is at market swap rates.[7] Mathematically,

[6]This discussion abstracts from haircuts and collateral agreements on both the repo and swap legs of the asset swap.

[7]This condition implicitly assumes that collateral arrangements eliminate any counterparty risk from the swap. See Chapter 13.

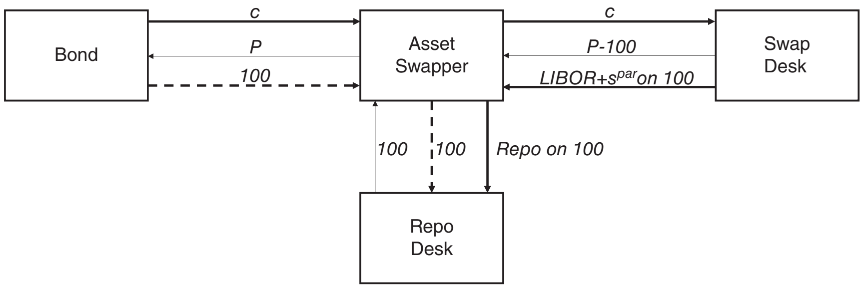

FIGURE 14.2 A Par-Par Asset Swap with Financing.

let A^{fixed} and $A^{floating}$ be the factors such that cA^{fixed} gives the present value of the payments of c on the fixed side of the swap and such that $s^{par}A^{floating}$ gives the present value of the payments of s^{par} on the floating side.[8] Let d be the discount factor for cash flows at maturity. Also, along the lines of Chapter 13, include a fictional notional amount of 100 at maturity on both legs of the swap, and note that the present value of receiving LIBOR and the final notional amount is par. Then, the fair pricing condition for the swap is,

$$(P-100)+100+100s^{par}A^{float} = cA^{fixed}+100d$$

$$s^{par} = \frac{cA^{fixed}+100d-P}{100A^{float}} \tag{14.1}$$

From Equation (14.1), for a given swap rate curve, as the credit risk of the bond increases, P decreases, which, in turn, increases the asset swap spread, s^{par}. Table 14.7 reports that, as of August 15, 2021, the par asset swap spread of the Genworth 4.90s of 08/15/2023 is 526.4 basis points.

A second flavor of asset swaps, namely, a *market value asset swap*, is illustrated in Figure 14.3. The market value asset swap differs from the par asset swap in that P is borrowed from the repo desk, rather than 100; $P-100$ is paid by the swap desk at the maturity rather than the initiation of the swap; and, as a result, LIBOR plus the spread, s^{mkt}, is earned on P rather than 100. Following the same notation and logic as in the case of the par asset swap, the fair market value asset swap spread is given by,

$$P+s^{mkt}PA^{float}+(P-100)d = cA^{fixed}+Pd$$

$$s^{mkt} = \frac{cA^{fixed}+100d-P}{PA^{float}}$$

$$= \frac{100s^{par}}{P} \tag{14.2}$$

[8] A^{fixed} and $A^{floating}$ do not equal each other when the frequency of payments on the fixed and floating sides differ.

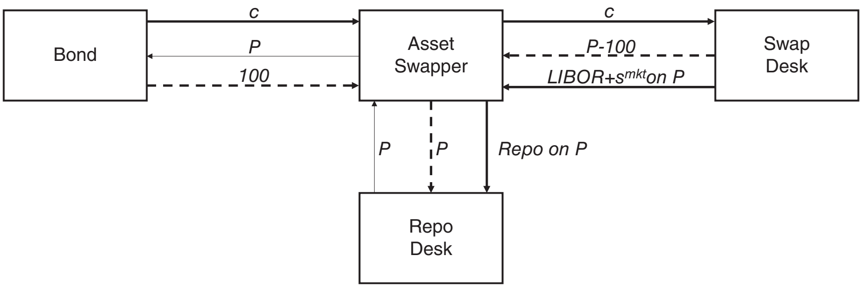

FIGURE 14.3 A Market Value Asset Swap with Financing.

where the final equality uses Equation (14.1). According to Table 14.7, the market value asset swap spread of the Genworth 4.90s of 08/15/2023, as of August 15, 2021, is 533.3 basis points. The relationship between the two asset swap spreads is quite intuitive: an investor can earn s^{par} on 100 or s^{mkt} on P, which, from Equation (14.2), gives the same result. The choice between the two asset swap trades, therefore, depends not on earnings, but on collateral and counterparty risk considerations.[9]

The final measure of credit spread considered in this section is an *effective spread* for corporate FRNs no longer priced at par. When an FRN is priced at par, its spread is particularly easy to interpret as a spread over the short-term rate benchmark received in exchange for bearing credit risk. Over time, however, as the credit quality of the issuer changes, the price of an FRN with a fixed spread changes. As a result, because an investor pays a premium or gets a discount to face amount when buying the FRN, in addition to its spread, that spread is no longer as easily interpreted.

The effective spread converts an FRN's premium or discount into a run rate and adds it to the actual spread. Let the actual spread be s^{float}; the price of the floater P; the present value factor, as before A^{float}; and the effective spread s^{eff}. Then, investors are indifferent between receiving s^{float} for a price of P and s^{eff} for a price of 100 if,

$$100s^{eff}A^{float} - 100 = 100s^{float}A^{float} - P$$

$$s^{eff} = s^{float} + \frac{100 - P}{100A^{float}} \tag{14.3}$$

[9]Begin with the case of a premium bond, that is, $P > 100$. In a par asset swap, the swap desk advances money to the asset swapper, which typically requires the asset swapper to post collateral. Over time, however, as the asset swapper makes coupon payments, the counterparty risk and collateral requirements decline. In the market value asset swap, there is no initial swap payment and, therefore, no initial collateral requirements. Over time, however, as the asset swapper makes payments, the obligations of the swap desk to pay $P - 100$ at maturity gives rise to collateral obligations from the swap desk to the asset swapper. In the case of a discount bond, that is, $P < 100$, the collateral implications are reversed.

This section concludes by noting that the trades described here for earning the credit spread are subject not only to the risk of default, but also to financing risk. Whenever a bond is purchased for a relatively long-term holding period, but is financed with short-term repo, there is the risk that the repo rate will increase by more than the discounting or benchmark rates. In that case, the bond will earn less than the benchmark short-term rate (e.g., LIBOR) plus the spread. An even more extreme risk is that repo lenders will refuse to roll positions, because they no longer wish to lend to the borrower, because they no longer wish to lend against a particular bond, or because they need the cash themselves. In that scenario, corporate bond investors with short-term repo financing will have to sell their bonds, most likely at a loss, to repay outstanding repo loans. Financing risk in the context of credit risk appears again, later in this chapter, as a key difference between bonds and credit default swaps.

14.4 CREDIT RISK PREMIUM

If credit spreads, on average, just compensated for credit losses, then risk-averse investors would just as soon buy safe government bonds at the same average return, but without the downside risk. For corporate debt to be attractive, therefore, spreads must not only compensate for credit losses, but also provide an additional *credit risk premium*. This section reviews some evidence that, on average, corporate bond spreads do, indeed, more than compensate for credit losses.

One study estimates that the average corporate yield spread over default-free bonds in the United States from 1866 to 2008 was 153 basis points, while the average credit loss was only 75 basis points. This long-term evidence indicates a substantial credit risk premium, in that average spreads are about twice average credit losses.[10]

Another study, over the more recent period 2002–2015, also finds a significant credit risk premium. Defining the premium as credit spread minus expected losses, the second column of Table 14.8 reports the median ratio of premium to credit spread, and the third column reports the median ratio of premium to expected loss, by rating. To interpret these numbers, consider a Baa-rated bond at a spread of 200 basis points. At the premium to spread ratio of 76% given in the second column, 76% of 200 or 152 basis points are due to a risk premium, while only the remaining 48 basis points are

[10]Giesecke, K., Longstaff, F., Schaefer, S., and Strebulaev, I. (2011), "Corporate Bond Default Risk: A 150-Year Perspective," *Journal of Financial Economics* 102, pp. 233–250. The credit loss estimate comes from a 1.5% calculated default rate and an assumed 50% recovery rate.

TABLE 14.8 Median Ratios of Premium to Credit Spread and to Expected Loss by Rating, 2002–2015.

Rating	Premium/Spread (%)	Premium/Expected Loss
Aaa	59	1.43
Aa	65	1.83
A	68	2.11
Baa	76	3.12
Ba	80	4.01
B	77	3.37
Caa	71	2.49
Ca-C	68	2.12

Source: Berndt, Douglas, Duffie, and Ferguson (2018).

compensation for expected losses. Expressed in terms of the third column of the table, these 152 basis points of credit risk premium are over three times the expected loss of 48 basis points. Over the sample period of this study, bonds with Baa to B ratings are an investor sweet spot in the sense of providing the highest premium for the amount of risk borne.[11]

14.5 CREDIT DEFAULT SWAPS

In a *single-name credit default swap* (CDS), a *protection buyer* or *CDS buyer* pays a *fee*, *premium*, or *coupon* to a *protection seller* or *CDS seller* in exchange for a *compensation payment* in the event that an issuer defaults. Bonds and CDS together comprise the markets for trading credit risk.

A CDS contract is defined by a *reference entity*, a list of *credit events*, a term or maturity, *reference obligations*, and a notional amount. As an example, consider a five-year CDS on a notional amount of $1 million of Genworth Senior Unsecured 6.5s of 06/15/2034. A payment by the seller of this CDS would be triggered if, before its maturity, Genworth is determined by an industry-led *determinations committee* to have experienced a credit event, like a bankruptcy filing or a failure to pay outstanding obligations.[12] In that scenario, the protection seller must make the protection buyer whole with respect to $1 million face amount of the underlying bonds.

[11]Berndt, A., Douglas, R., Duffie, D., and Ferguson, M. (2018), "Corporate Credit Risk Premia," *Review of Finance*, pp. 419–454. The credit spreads in this study are measured using CDS spreads.

[12]Credit events for North American CDS typically do not include restructurings, which are usually part of the bankruptcy process. Restructurings are significant, however, to the determination of credit events in European CDS.

More specifically, through *physical settlement* of the contract, the protection buyer delivers $1 million face amount of the bond to the protection seller in exchange for $1 million. Alternatively, through *cash settlement* of the contract, if the price of the bond were determined to be, for example, $400,000, the protection seller pays $600,000 to the protection buyer, again making the protection buyer whole on the $1 million face amount. (The existence of several reference obligations and the CDS delivery option, along with the auction mechanism by which protection sellers compensate protection buyers, is discussed presently.)

In exchange for the compensation payment in the event of default, the protection buyer pays a quarterly premium until the earlier of the maturity of the CDS or the event of default, in addition to paying or receiving an *upfront amount* at the initiation of the contract. The details of the premium and the upfront amount are discussed herein. For now, however, note that the *CDS spread* refers to the actual or hypothetical annualized premium on a CDS with an upfront payment of zero. In other words, a protection buyer can be thought of as paying the CDS spread in exchange for compensation in the event of default.

Table 14.9 gives selected sovereign and corporate five-year CDS spreads as of November 2021. For example, the cost of compensation for a credit event in Greece over the subsequent five years is 112.3 basis points or 1.123% annually. (For sovereigns, which do not file for bankruptcy, credit events include a moratorium on or a repudiation of debt obligations in addition to a failure to pay.) The CDS spread for Genworth, at 378.3 basis points, implies a cost of $37,830 per year to insure against a default of the $1 million face amount of the 6.5s of 06/15/2034 mentioned already. The cost of insuring $1 million face amount of bonds for five years ranges in the table from a low of $910 per year for German government bonds to a high of $62,450 per year for the bonds of MBIA Insurance.

TABLE 14.9 Selected Sovereign and Corporate Five-Year CDS Spreads, as of November 2021. Spreads Are in Basis Points.

Sovereign	Spread	Corporate	Spread
Germany	9.1	Marsh & McLennan	25.3
United States	16.0	JPMorgan Chase	53.8
Spain	35.2	BBVA Bancomer	70.3
China	48.9	Ally Financial	105.7
Italy	89.6	Fairfax Financial	154.5
Greece	112.3	Banco do Brasil	218.5
Brazil	265.8	Genworth	378.3
Turkey	469.6	MBIA Insurance	624.5

Whereas single-name CDS allow market participants to trade the credit of a single name, *index CDS* allow for the trading of broader credit portfolios. The five most popular index CDS are: CDX.NA.IG, which represents a portfolio of 125 single-name CDS on North American (NA), investment-grade (IG) names; CDX.NA.HY, on 100 single-name CDS on North American, high-yield (HY) names; iTraXX Europe Main, on 125 single-name CDS on European, investment-grade names; iTraXX Europe Crossover, on 50 single-name CDS on European, high-yield names; and iTraxx Europe Senior Financials, on senior debt of 25 names in that sector. The detailed workings of CDS indexes are described in the London Whale case study presented later. Note that industry jargon is different for index than for single-name CDS. The buyer of index CDS **receives** the premium and **pays** compensation in the event of default – just like the buyer of a portfolio of bonds – while the seller of index CDS **pays** the premium and **receives** compensation in the event of default – just like the short seller of a portfolio of bonds.

Figure 14.4 graphs historical CDS spreads for the 10-year iTraxx Europe Main and iTraxx Europe Crossover (XO) indexes. The cost of protection for the XO index spiked during the financial crisis of 2007–2009, during the height of the European sovereign debt crisis in 2011–2012, and at the start of the pandemic and economic shutdowns in early 2020. (The behavior of CDX.NA.HY, not shown, was qualitatively similar, though it spiked to significantly higher levels during the financial crisis.) The fluctuations of the iTraxx Europe Main have been comparatively modest, reaching 200 basis points during the sovereign debt crisis, but settling in recent years at about 100 basis points.

Figure 14.5 shows the slope of the term structure of CDS spreads for the iTraxx Europe XO index, in particular, the difference between the 10- and three-year index CDS spreads. In normal times, the annual cost of protection is higher for 10 years than for three years, because greater possibilities of disruption arise in the more distant future. During the financial crisis and at the height of the European sovereign debt crisis, however, the cost of three-year protection was greater than or equal to the cost of 10-year protection. In particularly stressful financial times, near-term events are the most uncertain, after which – at least for all surviving entities – credit risk is likely to have returned to levels that prevail in calmer times.

Uses of CDS

This introduction to CDS now turns to market participants. An often cited use case is a corporate bond investor hedging the default risk of a bond by buying CDS protection. This case does not really stand on its own, however, because an investor is likely to give up most if not all of the bond's credit spread by buying protection. Or, put another way, an investor not wanting

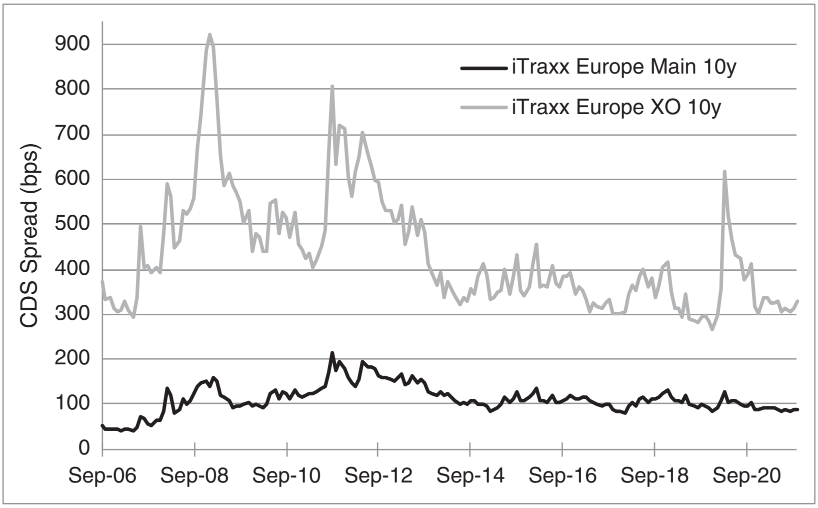

FIGURE 14.4 iTraxx Europe 10-Year CDS Indexes.

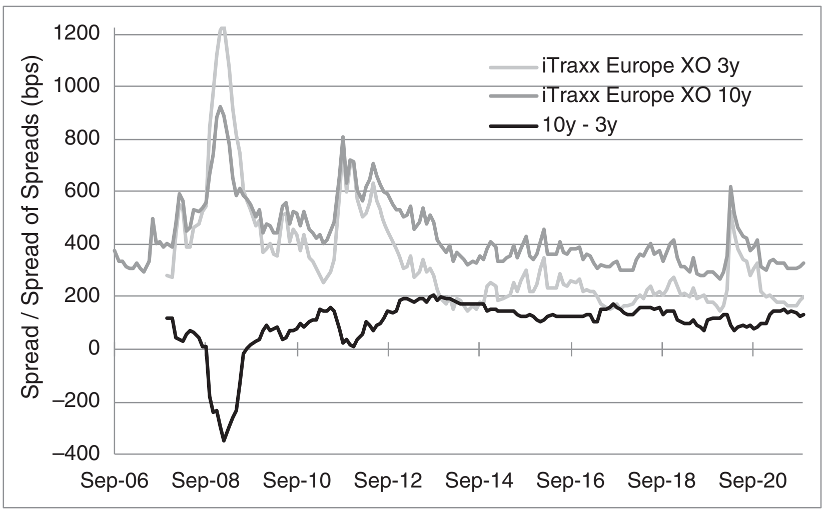

FIGURE 14.5 iTraxx Europe 10-Year versus Three-Year Crossover CDS Indexes.

to bear the credit risk of a particular bond can just sell that bond. Other use cases, therefore, are worth exploring.

As mentioned earlier and discussed presently, selling CDS protection very much resembles buying a bond: both the protection seller and the investor receive coupons over time and lose money in the event of a default. Many sellers of protection, therefore, choose to bear credit risk through

CDS rather than bonds. One reason might be that the relatively few CDS contracts on a particular name (e.g., a five- and 10-year) are more liquid than that name's many outstanding bond issues. Similarly, selling protection on an index might be a lot less costly than assembling a similarly diversified portfolio by buying individual corporate bonds. Even if an asset manager ultimately wants to hold the bonds, the fastest way to put on the position might be to sell protection and then, over time, buy the bonds and unwind the protection sold. A second reason to prefer selling protection over buying bonds might be to isolate credit risk from interest rate risk. Many insurance companies, for example, engage in *replication synthetic asset transactions* (RSATs), in which, for example, they buy government bonds of a maturity that suits their asset–liability management requirements and then sell protection on a credit that they like for investment purposes. A third reason, also discussed in this chapter, is that selling protection allows an investor to take risk with relatively little capital, whereas buying a bond requires that cash be available or explicitly borrowed.

Protection buyers, on the other hand, very much resemble short sellers: both pay coupons over time and win money in the event of a default. One reason to short corporate credit in CDS form is that it may be most efficient to do so, given the relatively significant trouble and expense of borrowing corporate bonds to short them. A second reason, paralleling a discussion in the previous paragraph, is that an investor might want to sell a corporate bond, but its liquidity would require a gradual, relatively slow unwind. In that case, if urgency is important, the best strategy might be to buy protection immediately in the more liquid CDS market, and then gradually sell the bond and unwind the purchased protection. A third reason to buy CDS protection is to hedge against risks that are not traded at all. For example, an entity doing business with Italian firms that have no publicly traded debt or related CDS might want to hedge against deteriorating credit conditions in that country by buying protection on Italian banks. This example has important consequences, as it implies that market participants might purchase protection on credits to which they have no direct exposure. And this, in turn, means that the amount of CDS outstanding on particular credits might exceed their outstanding debt. Back to the example, if many market participants want to hedge financial risk in Italy, the amount of CDS protection bought on Italian banks could exceed the amount of Italian bank debt outstanding.

A recent study puts the global outstanding notional amount of CDS at \$9.4 trillion, split 62% in index CDS and 38% in single-name CDS. Notional amounts are misleading indicators of outstanding risk, however, for the same reasons that apply to interest rate swaps, which are discussed in Chapter 13. Net notional outstanding then, from the same study, which nets each entity's long and short positions, is only \$1.5 trillion, with two-thirds of that in index CDS and one-third in single-name CDS. CDX.NA.IG and iTraxx

Europe account for about 70% of traded index amounts, and the top five aforementioned CDS contracts account for more than 90% of index activity. With respect to the single-name market, about 800 distinct names are referenced in each quarter of trading, but only about 550 of those names are referenced in every quarter. This implies that the single-name market consists of a core group of names that are traded consistently over time and other names that trade when interest is high in their particular credit situations.[13]

14.6 CDS UPFRONT AMOUNTS

Before the financial crisis of 2007–2009, CDS traded like interest rate swaps; that is, the coupon changed every moment with market conditions, and standard maturities were fixed terms from the settlement date. For example, buying $1 million notional of five-year CDS on Genworth at a spread of 558.92 basis points on August 16, 2006, was a commitment to pay $55,892 annually in exchange for compensation in the event of a Genworth default through August 16, 2011. Consequently, as is still true for interest rate swaps today (see Chapter 13), unwinding CDS trades was relatively difficult. For example, to unwind the CDS just described after one month, a trader would most like to sell protection at a fee of 558.92 basis points through August 16, 2011. But at the time of the unwind, on September 16, 2006, the most liquid or on-the-run five-year CDS would mature on September 16, 2011, and might carry a spread of 490 basis points. The trader, therefore, would have to incur relatively high transaction costs to unwind the existing CDS, or would have to sell protection through the on-the-run CDS and manage the maturity and spread mismatch between the original and the hedging CDS.

Since the financial crisis, as part of a broader push by regulators to improve operations in this market, CDS have become more standardized. First, maturity dates are limited to *pseudo-IMM* (international money market) dates, that is, the 20th days of March, June, September, or December. Thus, for example, all five-year contracts traded between June 21, 2021, and September 20, 2021, mature on September 20, 2021; all five-year contracts traded between September 21, 2021, and December 20, 2021, mature on December 20, 2021; etc. Second, all contracts have annual coupons of either 100 or 500 basis points, with a market determined *upfront amount* to compensate for the difference between the market CDS spread and the standardized coupon. Upfront amounts are described in greater detail later, but, for the present, note how these two changes in market conventions simplify the unwind of CDS. Buying a $1 million five-year CDS on Genworth

[13] ISDA (2019), "Global Credit Default Swaps Market Study," September. The sample period is from January 2014 to June 2019.

on August 16, 2021, at a spread of 558.92 actually requires making annual payments of 500 basis points or \$50,000, and – as shown next – an upfront payment of \$23,190. Furthermore, the contract matures on September 20, 2021. Therefore, to unwind this contract after one month, the trader can sell a still on-the-run CDS, that is, a CDS with a coupon of 500 basis points that matures on September 20, 2021, which exactly offsets the future cash flows of the original CDS.[14] If the market upfront payment fell over the month to \$13,000, then the trader suffers a loss of \$10,190 relative to the original purchase price. Unwinding the contract after more time has passed can be more difficult, because the on-the-run contract may have changed, but some liquidity is likely still available: many market participants likely traded Genworth CDS between June 21, 2021, and September 20, 2021, and, therefore, have contracts with a coupon of 500 basis points and a maturity of September 20, 2021.

A short mathematical prelude is needed here before moving to the calculation of upfront amounts. Simple credit risk models assume a constant *hazard rate*, λ, defined such that the probability of default over a short time interval, dt years, equals λdt. If the hazard rate is 10% and that time interval is six months, then the probability of default over the six months is 10%/2 or 5%, and the probability of "survival," meaning no default, is 95%. To survive over a year requires survival over both the first and second six months, which has a probability of $95\% \times 95\% = 90.25\%$. And this survival probability, in turn, implies that the probability of default over the year, whether in the first or the second six months, is 1 – 90.25%, or 9.75%. Note that breaking the one-year time frame into two six-month periods makes the probability of survival over the year, 90.25%, greater than 90% (100% minus the annual hazard rate of 10%): with 95% survival over the first six months, only that 95% – not the full 100% – is subject to default over the second six months. Correspondingly, the probability of default over the year, 9.75%, is less than the annual hazard rate of 10%.

More generally, Appendix A14.1 shows that if the hazard rate is applied continuously, then the *cumulative survival probability* over T years, $CS(T)$, and the *cumulative default probability* over T years, $CD(T)$, are,

$$CS(T) = e^{-\lambda T} \tag{14.4}$$

$$CD(T) = 1 - e^{-\lambda T} \tag{14.5}$$

[14]If the CDS is cleared, as most are today, then the buy and sell described in the text actually cancel, leaving no position with the central counterparty. If the trades are not cleared, and the buy is with a different counterparty than the sell, then the promised cash flows cancel, but the two contracts remain in force and are subject to counterparty risk.

TABLE 14.10 Calculating the Upfront Amount for 100 Notional Amount of an Annual Pay Five-Year CDS on Genworth at a Coupon of 500 Basis Points and a CDS Spread of 558.92 Basis Points, as of August 16, 2021. The Recovery Rate Is 40%. Probabilities and the Hazard Rate Are in Percent.

Year	Discount Factor	Cumulative Survival Probability	Default Probability
1	0.998462	91.099	8.901
2	0.994257	82.991	8.109
3	0.985062	75.604	7.387
4	0.973167	68.875	6.729
5	0.959797	62.744	6.130
	Hazard Rate		9.322
	Value of Fee Leg		21.995
	Value of Contingent Leg		21.995
	Upfront Amount		2.319

To illustrate the use of these equations with the same annual hazard rate of 10%, the cumulative survival probability over four years is $e^{-10\%\times 4} = 67.0\%$, and the cumulative default probability is one minus that, or 33.0%. Note again that, because of the compounding of survival probabilities over many short time intervals, the four-year survival rate of 67.0% is significantly greater than one minus four years at 10%, or 60%, and, correspondingly, the four-year default rate of 33.0% is significantly less than 40%.

Appendix A14.2 gives general, algebraic formulae for the market convention of calculating a CDS upfront amount given a quoted CDS spread. The text, however, continues with an example described in Table 14.10: 100 notional amount of a five-year Genworth CDS as of August 16, 2021. The CDS coupon is 500 basis points; the market CDS spread is 558.92 basis points; and the assumed recovery rate is 40%.[15] For simplicity, it is assumed that premium payments are annual (instead of quarterly).

The basic idea behind the market convention is as follows. First, find the hazard rate that makes the Genworth CDS "fair" in the sense that the expected discounted value of paying for protection with the CDS spread – here 558.92 basis points per year – is equal to the expected discounted value of receiving compensation in the event of default – here $100 \times (1 - 40\%) = 60$. This hazard rate turns out to be 9.322%. Second,

[15]Market conventions assume a recovery rate of 40% for many, but not all, CDS contracts. For example, contracts on North American subordinated debt assume a recovery rate of 20%, and contracts on European emerging market debt assume a recovery of 25%.

using that hazard rate, find the expected discounted value to the protection buyer of paying the actual CDS coupon – here 500 basis points – instead of the CDS spread of 558.92 basis points. That value turns out to be 2.319. Hence, to buy 100 face amount of Genworth protection at a coupon of 500 basis points, when the CDS spread is actually 558.92, the protection buyer has to pay an additional 2.319 upfront.

The first column of Table 14.10 lists the years of the five annual payments. The second column gives the discount factors from the LIBOR swap curve as of the pricing date. The third and fourth columns give the cumulative survival and default probabilities at the end of each year, computed with Equations (14.4) and (14.5) and a hazard rate of 9.322%. This hazard rate is shown in the table, and its derivation is given presently.

Like all CDS, the Genworth CDS in the example can be described as having two legs: a *fee leg* and a *contingent leg*. The fee leg comprises the payments of the premium from the buyer to the seller of protection. At the end of any year in which there has not been a default, the buyer in the current calculation has to pay the CDS spread of 558.92 basis points on the 100 notional amount, or 5.5892. If there is a default over the year, then the buyer has to pay the accrued coupon from the start of the year to the time of default. If, for example, the default occurred after three months, or one-quarter of a year, the buyer would have to pay one-fourth of the coupon, or 1.3973; if the default occurred in the middle of the year, the buyer would have to pay 2.7946. For simplicity, however, the market convention for calculating upfront amounts assumes accrual for half the period, or, in this example, for half of the year. (This assumption has less impact on the calculations in practice, because coupons are actually paid quarterly.)

Continuing with the example, then, the buyer's payment on the fee leg at the end of the first year is 5.5892 if there is no default over the year, and 2.7946 if there is. The probability of these two contingencies are reported in the table as 91.099% and 8.901%, respectively. Hence, the expected present value of this payment is,

$$0.998462[91.099\% \times 5.5892 + 8.901\% \times 2.7946] = 5.332 \qquad (14.6)$$

Analogous calculations can be repeated for each of the next four years, using the appropriate survival and default probabilities and discount factors. Then, summing the results gives the present value of the fee leg, reported in Table 14.10, as 21.995.

With respect to the contingent leg, the seller of protection pays the buyer 60 in any year that Genworth experiences a default. The expected discounted value of that contingent payment in the third year, for example, is,

$$0.985062 \times 7.387\% \times 60 = 4.366 \qquad (14.7)$$

Performing this calculation for each of the five years and summing the results gives the present value of the contingent leg, reported in Table 14.10, as 21.995.

Because the discounted expected values of the fee and contingent legs are equal, the hazard rate of 9.322% correctly reflects the market CDS spread of 558.92 basis points. Solving for this hazard rate in the first place requires iterating through the calculations just described. For example, after setting up the calculations in Table 14.10 as an Excel spreadsheet, the built-in solver can be used to find the hazard rate that results in equal values for the fee and contingent legs.

The stage is now set for calculating the upfront amount such that paying this upfront amount at the initiation of the CDS and then paying a running premium of 500 basis points has the same discounted expected value as paying nothing upfront and then a running premium of 558.92 basis points. As derived previously, the expected discounted value of the fee leg making payments of 5.5892 is 21.995. It follows, then, that the value of payments of the standardized coupon of five in the fee leg is $21.995 \times 5/5.5892 = 19.676$. Therefore, the required upfront payment is just the difference, $21.995 - 19.676$, or 2.319: an upfront payment of 2.319 and a running premium of five (which is worth 19.676) has the same value as no upfront payment and a running premium of 5.5892 (which is worth 21.995).

A moment's reflection reveals that the upfront amount is positive whenever the CDS spread is greater than the standardized coupon. If the market CDS spread of Genworth were 450 on August 16, 2021, then the upfront amount turns out to be −2.051. Buying protection by *receiving* 2.051 upfront and paying a running premium of five has the same value as no upfront amount and paying a running premium of 4.50.[16]

This section concludes by emphasizing that the quoting of an upfront amount from a CDS spread is only a convention, like the relationship between price and yield-to-maturity. Market participants form their own views on upfront amounts at which they are willing to trade CDS. They may not be willing, for example, to assume a particular recovery rate or that the hazard rate is constant over time. But they all use the accepted market conventions to calculate upfront amounts from a quoted CDS spread or *vice versa*.

---

[16]As an exercise, the reader might want to verify in a table analogous to Table 14.10 that, at a CDS spread of 450, the hazard rate is 7.504%; the expected discounted value of each leg is 18.462; and therefore, as reported in the text, the upfront amount per 100 face amount is −2.051.

14.7 CDS-EQUIVALENT BOND SPREAD

The credit spreads defined earlier in the chapter are all measures of bond return assuming no default. An alternative approach, the *CDS-equivalent bond spread*, accounts for default and recovery and is computed along the lines of the previous section. The basic idea is to find the hazard rate such that the market price of the bond equals the expected discounted value of its cash flows. Then, the bond's CDS-equivalent spread is the CDS spread corresponding to that hazard rate. To illustrate, say that the market price of a five-year, 5% (annual pay) bond on Genworth is 94.561 as of August 16, 2021. It is shown next that the expected discounted value of that bond's cash flows equals that market price if the hazard rate is 9.322%. Furthermore, from Table 14.10, the five-year CDS spread at a hazard rate of 9.322% is 558.92 basis points. Therefore, the CDS-equivalent spread of this five-year, 5% bond is 558.92 basis points. In other words, the credit risk implied by the bond's market price is the same as that implied by a five-year CDS at a spread of 558.92 basis points. This spread depends on the restrictive assumption that the hazard rate is constant. The calculation of expected discounted values, while seeming to imply risk neutrality, is not as restrictive as it seems: the hazard rate can be considered "risk-neutral" so that it prices fixed income securities without necessarily reflecting real-world probabilities. (This distinction is discussed in Chapter 7.)

Appendix A14.4 gives the algebra for calculating the expected discount value of a bond's cash flows given a hazard rate. The text, however, continues with pricing 100 face amount of an annual pay, five-year, 5% Genworth bond as of August 16, 2021, and the results are summarized in Table 14.11. Because the bond's assumed market price is consistent with a hazard rate of 9.332%, which is the hazard rate in Table 14.10, the cumulative survival and default probabilities from that table can be used here. Also, because the pricing dates are the same, the discount factors from that table can be used here as well.

Along the lines of the previous section, the expected value of a bond's coupon in a given year is half of its value times the probability of default over the year plus its value times the probability of no default over the year. The discounted expected value of the first coupon of the 5% Genworth bond, for example, is as in Equation (14.6), but with the bond coupon of 5.00 replacing the CDS spread payment of 5.5892,

$$\begin{aligned}\text{Expected PV} &= 0.998462[8.901\% \times 2.50 + 91.099\% \times 5.00] \\ &= 0.222 + 4.548 = 4.770 \qquad (14.8)\end{aligned}$$

The two components of this coupon's expected discounted value are given in the second and third columns of the first row of Table 14.11. The next

TABLE 14.11 Calculating the Expected Discounted Value of 100 Face Amount of an Annual Pay Five-Year 5% Genworth Bond as of August 16, 2021, at a Hazard Rate of 9.322% and a Recovery Rate of 40%.

	Expected PV Coupon		Expected PV Principal	
Year	Default	No Default	Default	No Default
1	0.222	4.548	3.555	
2	0.202	4.126	3.225	
3	0.182	3.724	2.911	
4	0.164	3.351	2.620	
5	0.147	3.011	2.354	60.222
Total	0.916	18.760	14.663	60.222
Bond Price				94.561

rows of the table repeat this calculation for subsequent coupon payments, using the appropriate discount factors, cumulative survival probabilities, and cumulative default probabilities from Table 14.10. Summing the results, the total discounted expected value of the bond's coupons is $0.916 + 18.760 = 19.676$.

With respect to principal, if there is a default in any year, then the principal received is the face amount times the recovery rate. The discounted expected value of principal received in the third year, for example, is the three-year discount factor times the probability of default in year three times the recovery on 100 face amount,

$$0.985062 \times 7.3787\% \times 40 = 2.911 \tag{14.9}$$

which is the value given in the fourth column, third row, of Table 14.11. (Note that, in Equation (14.9), the bondholder receives 40, while in Equation (14.7), the buyer of protection receives a default compensation of 60.) The table uses the equivalents of Equation (14.9) to calculate the rest of the fourth column.

If there is no default up to and including maturity, the bond pays its full principal amount. The expected discounted value of this cash flow is,

$$0.959797 \times 62.744\% \times 100 = 60.222 \tag{14.10}$$

and is given in the year five row and in the rightmost column of Table 14.11.

Finally, adding together the four components of the bond's expected discounted value from the table, the bond's price is 94.561. Hence, because the bond is fairly priced using a hazard rate of 9.322%, and because that hazard rate corresponds to a CDS spread of 558.92 basis points (from the previous section), this bond's CDS equivalent spread is 558.92 basis points.

This section now compares the CDS-equivalent bond spread just explained with the bond spread discussed already. First, Appendix A14.3 shows that the CDS spread is approximately equal to the hazard rate times one minus the recovery rate. Mathematically, letting s^{CDS} denote the CDS spread; letting λ denote the hazard rate, as before; and letting R denote the recovery rate,

$$s^{CDS} \approx \lambda(1 - R) \tag{14.11}$$

In the context of Table 14.10, for example, the approximation predicts a CDS spread of 9.322% times (1 − 40%), or 559.32 basis points, while the actual spread is a very close 558.92 basis points.

Second, Appendix A14.5 shows that,

$$s^{Bond} \approx \lambda(1 - R^m) \tag{14.12}$$

where s is the bond spread, and R^m is the recovery rate as a percentage of market value. To this point the chapter has assumed *par recovery*, that is, bond recovery rates are best modeled as a fixed percentage of face amount. *Market recovery*, by contrast, which is used to derive Equation (14.12), assumes that recovery is a fixed percentage of market value. Say that two bonds were sold by the same issuer with the same seniority, but one of the bonds has a much longer maturity and trades at a larger price discount or premium, depending on the level of rates and spreads. In the event of default, will the two bonds recover the same amount – as assumed by par recovery – or will the longer-term bond's recovery reflect its greater discount or premium – as assumed by market recovery? Par recovery is the more common assumption and is better supported by empirical evidence.[17]

In short, CDS and bond spreads are not equivalent, and the consensus in favor of the par recovery assumptions argues in favor of preferring the CDS spread. Furthermore, at high hazard rates and low bond prices, the CDS spread is significantly greater than the bond spread. To illustrate, recall from Tables 14.10 and 14.11 that the five-year 5% Genworth bond, at a hazard rate of 9.322% and a price of 94.561, has a CDS spread of 559 basis points. From the data in these tables, it can also be computed that its bond spread is a similar 551 basis points.[18] However, if the hazard rate is a much higher

---

[17]See Guha, R., Sbuelz, A., and Tarelli, A. (2020), "Structural Recovery of Face Value at Default," *European Journal of Operational Research* 283, pp. 1148–1171; and Bakshi, G., Madan, D., and Zhang, F. (2006), "Understanding the Role of Recovery in Default Risk Models: Empirical Comparisons and Implied Recovery Rates," FDIC Center of Financial Research Working Paper, September 6.

[18]Use the discount factors in Table 14.10 to compute forward rates. Then, find the spread that, when added to these forward rates, recovers – by simple discounting – the market price of 94.561.

16.70%, then the bond price is 80.688; the CDS spread is 1,000 basis points; and the bond spread is a much lower 932 basis points. (These computations are also left as an exercise for the reader.) Intuitively, by assuming too low a recovery for discount bonds, bond spreads have to be lower to reproduce seemingly high market prices.

14.8 CDS-BOND BASIS

Credit risk today is traded in both the corporate debt market and in the CDS market. It is natural to ask, therefore, whether a particular credit trades at the same price in both markets or is cheaper in one market than the other. If the latter, then there might be relative value trading opportunities to buy in one market and sell in the other.

Table 14.12 sets the stage with a simplified relationship between selling CDS protection on a particular credit and buying a bond of that same credit, financed in the repo market.[19] The table is a simplification by assuming that i) the CDS contract has an upfront payment of zero, that is, the CDS coupon equals the CDS spread; ii) the corporate bond is priced at par; and iii) the CDS and the underlying bond mature on the same date. Under these assumptions, the table shows that selling CDS protection is equivalent to buying a bond and financing its purchase with term repo to the bond's maturity. More specifically, the cash flows from selling 100 notional of CDS are: zero today; the spread s on 100 until the earlier of default or maturity; $-100(1-R)$ if and when the bond defaults; and zero if the bond matures without a default. The cash flows from buying 100 face amount of the bond at a price of par are -100 today; the coupon rate c on the face amount to the earlier of default or maturity; $100R$ if and when the bond defaults; and

TABLE 14.12 A Simplified Arbitrage Relationship Between Selling Protection on a Credit and Buying a Bond on That Same Credit, Financed by Repo.

	Today	Interim (%)	Default	Maturity/No Default
Sell CDS Protection	0	s	$-100(1-R)$	0
Buy Par Bond	-100	c	$100R$	100
Sell Repo	100	$-r$	-100	-100
Total	0	$c-r$	$-100(1-R)$	0

[19]For a more general discussion relating derivative positions to financed cash positions, see Tuckman (2013), "Embedded Financing: the Unsung Virtue of Derivatives," *Journal of Derivatives*, Fall, pp. 73–82.

100 if the bond matures without a default. Finally, the cash flows from borrowing the bond's purchase price through repo are 100 today; $-r$ on 100 until the earlier of default or maturity; and -100 at the earlier of default or maturity. (To clarify these repo cash flows, note that, if a bond defaults, the borrower has to unwind the repo position by paying off the 100 loan amount, which then ends interest payments on the loan.) Finally, adding the cash flows from purchasing the bond and selling the repo gives the "Total" row, which exactly matches the cash flows of selling CDS protection so long as $s = c - r$.

The *CDS-bond basis* refers to the difference between the CDS spread and some measure of the bond's spread over riskless rates. In the simplified setting of Table 14.12, where the bond spread is the difference between the par coupon and the matching-term repo rate, the CDS-bond basis is $s - (c - r)$. If the basis is positive, then $s > c - r$, which means that the bond is rich, or that it has a low spread relative to the CDS spread. Put another way, selling CDS protection earns more than buying the bond. If the basis is negative, then $s < c - r$, which means that the bond is cheap relative to the CDS spread, and that buying the bond earns more than selling CDS protection. In either case, when the basis is not zero, an arbitrage opportunity is available. If positive, execute a *positive basis trade*, that is, sell protection, short the bond, and buy the repo, to lock in $s - (c - r) > 0$ per period until the earlier of default or maturity. And if the basis is negative, execute a *negative basis trade*, that is, buy protection, buy the bond, and sell the repo, to lock in $(c - r) - s > 0$ per period.

Crucial to these arbitrage arguments is that the term of the repo equal the maturity of the bond and the CDS. In a negative basis trade, for example, selling overnight – rather than term – repo exposes a trader to financing risk: the overnight repo rate might increase dramatically or repo funding might be withdrawn completely, either because the bond's creditworthiness has deteriorated, because general financing supply has tightened, or because the trader's own credit has come under stress. In any of these scenarios, the trader might very well be forced to unwind the position at a loss, that is, when the bond has become even cheaper relative to CDS. Similarly, in executing a positive basis trade with overnight repo, the financing risk is not being able to continue to borrow the bonds in order to maintain the short position.

In practice, because there is no market for very long-term corporate repo, it is nearly impossible to execute a basis trade without bearing financing risk. Therefore, despite arbitrage arguments like Table 14.12, there is a fundamental difference between CDS and levered bond positions: CDS positions have implicit financing to maturity, while levered bond positions are inherently subject to financing risk.

Outside the simplified setting of Table 14.12, market participants compute the CDS-bond basis using different measures of a bond's spread, like yield spread, bond spread, asset swap spread, and CDS-equivalent bond spread. The advantages and disadvantages of each of these is discussed previously, but all of these measures can be misleading in the context of the CDS-bond basis. Without a term-repo rate, any measure of the basis ignores the financing risk in a CDS versus levered bond arbitrage trade.

A lot of money was lost in negative basis trades through the financial crisis of 2007–2009. As funding conditions became more difficult, CDS-bond bases moved negative, reflecting the difficulty of going long credit risk with levered bond positions relative to selling CDS protections. Some traders, however, saw these negative bases as trading opportunities, which turned into nightmares as funding conditions deteriorated even further. Over the crisis, in fact, the investment-grade index CDS-bond basis fell from about zero to negative 250 basis points.

Figure 14.6 shows the CDS-bond basis for French and Italian government bonds from January 2006 to October 2021. The basis is defined here as the difference between the five-year CDS spread and the five-year government bond yield spread over Germany. By way of economic backdrop, the European sovereign debt crisis ran from the end of 2009 to 2012, during which time the finances of both banks and governments in the so-called "peripheral" countries (e.g., Greece, Italy, Portugal, and Spain) came under pressure. Later, in the fall 2018, European Union and Italian government officials squabbled over Italy's proposed budget, which did not comply with existing fiscal rules. The figure shows that the sovereign Italian CDS-bond basis became significantly negative during both of these stressful periods.

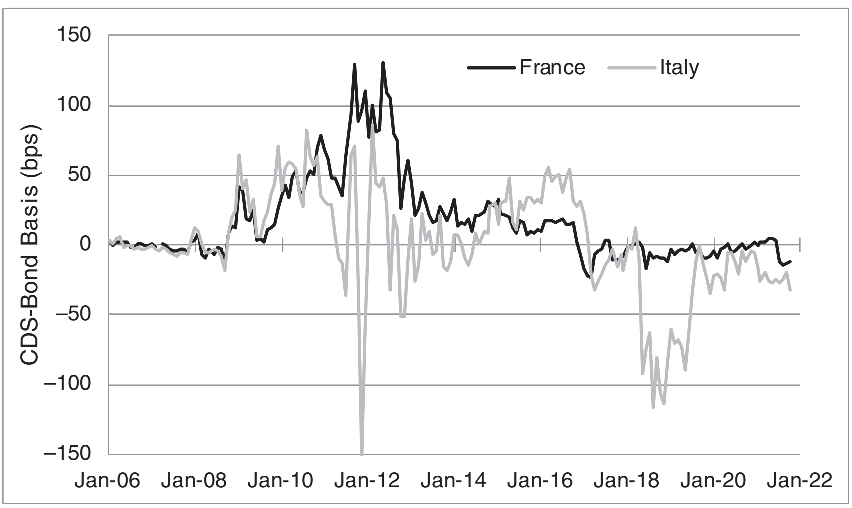

FIGURE 14.6 CDS-Bond Basis for French and Italian Sovereign Debt, January 2006 to October 2021.

As in the financial crisis of 2007–2009, repo lenders became reluctant to fund Italian government bonds, pushing up their spreads and pushing their CDS-bond basis into negative territory. In fact – though not shown in the figure – spreads of Italian to German government bonds rose to over 600 basis points during the sovereign debt crisis, and to nearly 300 basis points during the budget standoff. By contrast, French government credit exhibited a positive CDS-bond basis during the sovereign debt crisis. French over German spreads did rise to a peak of about 100 basis points at that time, but the bigger story was the difficulty of borrowing and shorting French government bonds, at least in part because the European Central Bank was aggressively buying European sovereign debt. In any case, the resulting upward pressure on French government bond prices resulted in a significantly positive CDS-bond basis.[20] Finally, more recently, funding stresses during the COVID pandemic and economic shutdowns drove both Italian and French government CDS-bond bases somewhat negative.

14.9 HAZARD-ADJUSTED DURATION AND DV01

The durations of bonds with credit risk are often calculated along the lines of Chapter 4, that is, cash flows are assumed to be paid on schedule, but are discounted at higher rates, typically at benchmark rates plus a credit spread or a term structure of credit spreads. Calculated this way, however, duration can be misleading for bonds with significant credit risk. For these bonds, expected cash flows are much earlier than scheduled cash flows and, consequently, their durations are correspondingly shorter.

The simple hazard-rate framework presented earlier can be used to calculate a *hazard-adjusted duration* (HAD) that accounts for the expected timing of cash flows given default. Along the lines of Table 14.11, given a benchmark rate curve and a recovery rate, find the hazard rate for which the expected discounted value of the bond's cash flows equals its market price. Then, shift the benchmark rate curve down and reprice the bond, using the same recovery and hazard rates. Finally, use the resulting price difference or percentage price difference to compute the hazard-adjusted DV01 or duration, respectively.

To illustrate the difference between the conventional and hazard-adjusted approaches, Figure 14.7 graphs conventional and hazard-adjusted durations for bonds of various terms at two different hazard rates. The benchmark

[20]For a rigorous discussion of funding and short-selling frictions during the European sovereign debt crisis, see Fontana, A., and Scheicher, M. (2016), "An Analysis of Euro Area Sovereign CDS and Their Relation with Government Bonds," *Journal of Banking and Finance* 62, pp. 126–140.

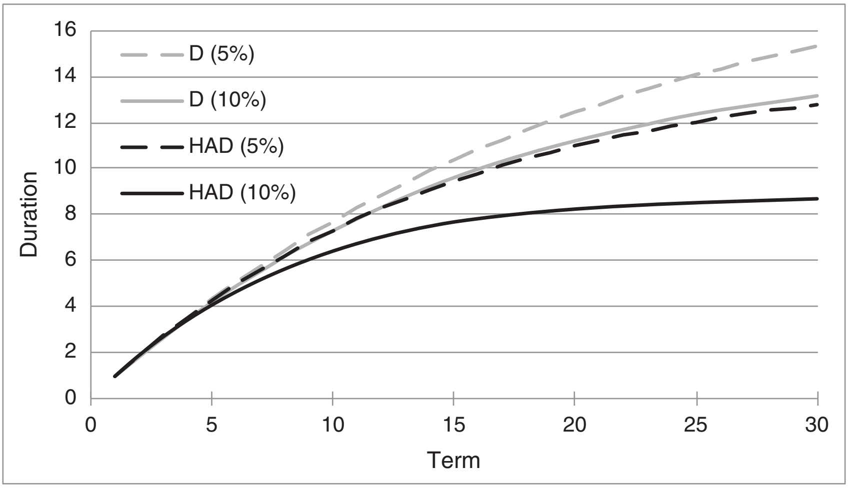

FIGURE 14.7 Conventional *Vs*. Hazard-Adjusted Durations.

rate curve is flat at 2%, and the recovery rate is fixed at 40%. The hazard rate is either 5% or 10%, which, given the 40% recovery rate, corresponds approximately to CDS spreads of 300 and 600 basis points, respectively. All bonds are priced using the constant hazard-rate model, and HADs are computed as described in the previous paragraph. Conventional durations are calculated by finding the fixed spread to the benchmark curve that correctly prices each bond, and then shifting the benchmark curve and repricing each bond, keeping its spread constant.

Bonds with a higher hazard rate have higher spreads and, therefore, lower durations than bonds with a lower hazard rate. The striking message of Figure 14.7, however, is that HADs can be significantly below conventional durations, particularly for larger hazard rates and longer terms. For 10-year bonds at a hazard rate of 5%, the conventional duration is 7.7 and the HAD is 7.3, while at a hazard rate of 10%, the durations are 7.3 conventional and 6.4 HAD. For 30-year bonds at a hazard rate of 5%, the conventional duration is 15.3 and the HAD is 12.8, while at a hazard rate of 10%, the durations are 13.2 conventional and 8.7 HAD.

14.10 SPREAD DURATION AND DV01

When trading or investing in credit portfolios, it is natural to measure the sensitivity of bond prices to changes in credit spreads. The standard measures, *spread duration* and *spread DV01*, shift the spread, keeping the benchmark curve and cash flows the same, calculate a new price, and then calculate a duration or DV01. Because rates used for discounting in this context are the sum of the benchmark rates and spreads, shifting the benchmark rate by one basis point results in the same price sensitivity as shifting the spread by one basis point, that is, spread durations and DV01s are the same as durations and DV01s with respect to interest rate changes. These spread sensitivities can still be useful, however, in a portfolio context. In a portfolio of government bonds, interest rates swaps, and corporate bonds, for example, spread DV01 with respect to swap spreads can be found by shifting spreads for swaps and corporate bonds only, while spread DV01 with respect to corporate spreads can be found by shifting spreads for corporate bonds only.

It is very common in the corporate bond setting to measure spread risk with *duration times spread*, or *DTS*, instead of spread duration. This methodology is based on the empirical regularity that changes in the spread of a corporate bond are proportional to the spread itself. Consider, for example, bonds A and B, which have spread durations of five and four years, and which trade at spreads of 100 and 250 basis points, respectively. Under the usual duration assumptions of parallel shifts, bond B has spread risk of 4/5 or 80% of that of bond A. But empirical evidence suggests that if the spread of bond A increases by 10%, that is, 10 basis points from 100 to 110, then the spread of bond B also increases by 10%, that is, 25 basis points from 250 to 275. In this case, after these spread changes, by the definition of duration, the percentage change of bond A is $5 \times 10/100 = 0.5$, and the percentage change of bond B is $4 \times 25/100 = 1$. Hence, bond B is actually twice as risky as bond A. This is reflected in their DTS: $5 \times 100 = 500$ for bond A, and $4 \times 250 = 1{,}000$ for bond B.

Price sensitivity to spreads can also be computed using the hazard-rate model described in this chapter. *Risky DV01* refers to the change in the value of a CDS contract for a one-basis-point change in the CDS spread. This can be computed, of course, by finding the hazard rate that changes the CDS spread by one basis point, recomputing the value of the CDS, and calculating the risky DV01. An equivalent bond metric can be computed similarly: find the hazard rate that changes the CDS-equivalent bond spread by one basis point, recompute its price, and calculate a DV01.

14.11 CDS SETTLEMENT AUCTIONS

Two challenges typical in the settlement of derivatives trades apply just as well to CDS. First, requiring that all protection buyers physically settle contracts leaves them open to a squeeze. As discussed earlier, protection buyers do not necessarily own the underlying bonds and the supply of bonds eligible for delivery might be subject to manipulation. Second, bond prices may not be readily observable for the purpose of cash settlement, because corporate bond liquidity – even if satisfactory in normal times – may be quite

TABLE 14.13 List of Deliverables for Hertz Corporation CDS Auction, June 23, 2020.

Description	Maturity
4.125% Sr. Notes	10/15/2021
7.625% Sr. Secured 2nd Priority Notes	06/01/2022
6.250% Sr. Notes	10/15/2022
5.500% Sr. Notes	03/30/2023
Letter of Credit Disbursements	12/18/2023
5.500% Sr. Notes	10/15/2024
7.125% Sr. Notes	08/01/2026
6.000% Sr. Notes	01/15/2028

limited through a default. Industry-run settlement auctions are designed to cope with both of these fundamental challenges.

Hertz filed for bankruptcy on May 22, 2020, and, on June 24, 2020, an auction was held to settle CDS referencing Hertz. Some physical settlement took place as part of the auction, and the final auction price of 26.375 was used for cash settlement. This section describes the CDS auction process in the context of this Hertz CDS auction.[21]

A single-name CDS contract specifies reference obligations of the reference entity, together with a set of rules, that determine whether an event of default has occurred and, if so, which bonds are deliverable, or eligible to be delivered, for the purposes of physical settlement. The final list of deliverable obligations into the Hertz CDS auction is given in Table 14.13.

The auction itself is divided into two phases. In phase one, participating dealers submit i) bids and offers for purchase and sale of deliverable obligations, with a predefined quotation size and maximum bid-offer spread; and ii) physical settlement requests. Table 14.14 lists the bids and offers of the dealers in the Hertz auction, with the bids sorted in descending order and the offers in ascending order. An *inside market midpoint* (IMM) is determined from these bids and offers as follows. First, discard all rows with *crossing* or *touching* markets (i.e., where the bid price exceeds or equals the offer price). In Table 14.14, the first four rows constitute crossing or touching markets, and are discarded. Second, using the top half of the remaining rows (rounded up, if necessary), compute the average of the bids and offers and round to the nearest eighth per 100 face amount. In Table 14.14, the bottom six rows remain. Therefore, using only the top three rows, the average of the bids and

[21]A more detailed description of the process can be found in Credit Suisse (2011), "A Guide to Credit Events and Auction," *Fixed Income Research*, January 12; and Markit (2010), "Credit Event Auction Primer," February.

TABLE 14.14 Dealer Initial Markets, Hertz Corporation CDS Auction, June 24, 2020.

Dealer	Bid	Offer	Dealer
Citigroup	26.0	22.5	Morgan Stanley
Credit Suisse	26.0	24.0	Barclays
Deutsche Bank	26.0	24.5	Goldman Sachs
RBC Capital Markets	25.5	25.5	Bank of America
BNP Paribas	**24.5**	**26.0**	J.P. Morgan Securities
J.P. Morgan Securities	**24.0**	**26.5**	BNP Paribas
Bank of America	**23.5**	**27.5**	RBC Capital Markets
Goldman Sachs	22.5	28.0	Citigroup
Barclays	22.0	28.0	Credit Suisse
Morgan Stanley	20.5	28.0	Deutsche Bank
Inside Market Midpoint	**25.375**		

offers, which are bolded, is 25.333, which, rounded to the nearest eighth, gives an IMM of 25.375.

Turning to the second part of phase one of the auction, Table 14.15 shows the physical settlement requests of the participating dealers. These requests must match the direction and size of their respective net CDS positions. Goldman Sachs and its clients, for example, by offering to deliver 140.0 million face amount of deliverable obligations, must, on net, have bought protection on at least 140.0 million CDS notional amount. Put another way, Goldman and its clients are requesting to collect on their having bought a net of 140.0 million notional amount of CDS by delivering 140.0 million face amount of deliverable obligations and receiving \$140 million. J.P. Morgan Securities and its clients, on the other side, by bidding to take 20.0 million of deliverable obligations, must, on net, have sold protection on at least 20.0 million CDS notional amount. They are requesting to settle their CDS obligations by taking delivery of 20.0 million of deliverable obligations and paying \$20 million. Taking the physical settlement requests all together, there are total requests to sell or deliver \$143.525 million face amount and total requests to buy or take delivery on \$42 million, leaving a net open interest (NOI) to sell of \$111.525 million.

Before proceeding to phase two of the auction, a note is made of *adjustment amounts*. To provide an incentive for dealers to submit competitive bids and offers, dealers are penalized for submitting off-market bids and offers, where "off-market" is judged relative to the IMM and NOI. More specifically, in the case of the Hertz auction, there was a NOI to sell and the IMM was 25.375. It follows from Table 14.14 that the bids and offers of Citigroup, Credit Suisse, and Deutsche Bank at 26.0–28.0, and those of RBC Capital Markets at 25.5–27.5, were too high in a market where there

TABLE 14.15 Physical Settlement Requests, Hertz Corporation CDS Auction, June 24, 2020. Amounts Are in $Millions.

Dealer	Size
Offers	
Bank of America	0.0
BNP Paribas	3.525
Citigroup	10.0
Deutsche Bank	0.0
Goldman Sachs	140.0
Morgan Stanley	0.0
RBC Capital Markets	0.0
Total Sell Requests	143.525
Bids	
Barclays	7.0
Credit Suisse	15.0
J.P. Morgan Securities	20.0
Total Buy Requests	42.0
Net Open Interest (Sell)	111.525

were more sellers than buyers. Therefore, these dealers had to pay adjustment amounts equal to the difference between their above-market bids and the IMM times the quotation amount: \$2 million × (26.0 – 25.375)/100 or \$12,500 for Citigroup, Credit Suisse, and Deutsche Bank, and \$2 million × (25.5 – 25.375)/100, or \$2,500 for RBC Capital Markets. In phase one auctions with NOI to buy, adjustment amounts are levied on dealers with bids and offers below the IMM.

Phase two of the auction takes place a few hours after the results of phase one. The purpose of phase two is to determine a final auction price, defined as the clearing price for the NOI established in phase one. In phase two, market participants – including, but not limited to the participating dealers in phase one – submit limit orders to buy, if the NOI was to sell, or limit orders to sell, if the NOI was to buy. Furthermore, to ensure the sensible result that the final auction price not be too high relative to the IMM if the NOI is to sell, or not be too low relative to the IMM if the NOI is to buy, limit order prices are constrained by a *cap amount*, usually set to half the maximum bid-offer spread in phase one. In the Hertz auction, the cap amount was half of two, or one. And, because there was a NOI to sell, the limit orders to buy in phase two were constrained to be less than or equal to the IMM of 25.375 plus the cap amount of one, or 26.375.

TABLE 14.16 Goldman Sachs Limit Orders, Hertz Corporation CDS Auction, June 24, 2020. Amounts Are $Millions.

Bid Price	Size
26.375	42
25.000	10
24.000	10
23.750	5
23.500	24
23.250	10
23.000	23
22.500	2
20.000	14

Table 14.16 shows the limit orders to buy submitted by Goldman Sachs in phase two of the Hertz auction. On behalf of itself and its customers, Goldman Sachs bid to purchase $42 million face amount at a price of 26.375; an additional $10 million at a price of 25.000; and so forth. Note that bids from the initial market in phase one of the auction carry over into this phase, so that Goldman Sachs' bid at 22.5 for the predefined size of $2 million in Table 14.14 appears as the penultimate row of Table 14.16.[22]

All of the limit orders to buy Hertz bonds are collected and ordered by price, from high to low. Then, following the rules of a *Dutch auction*, the final auction price is set such that the quantity of limit orders to buy at that price or above matches the $111.525 million NOI to sell. In the Hertz auction, at the highest allowable bid price of 26.375, there were limit orders to purchase $118.50 million face amount, which is more than the NOI available for sale. Therefore, the final auction price was set at 26.375 and all bidders at that price were allocated a pro rata amount of the available $111.525 million. Goldman Sachs, for example, with its limit order to buy $42 million at 26.375, was able to purchase $42 million × $111.525/$118.50, or $39.528 million face amount through the auction.

To summarize, the auction had two outcomes. First, $143.525 million face amount of deliverable Hertz bonds were sold by the dealers or their customers in the "Offers" part of Table 14.15 at a price of 26.375 to the dealers or their customers in the "Bids" part of the table – $42 million – and to the market participants awarded allocations in phase two of the auction – $111.525 million. Second, all Hertz CDS settled at the final auction price of 26.375. CDS counterparties who chose cash settlement

[22] When NOI is to sell, crossing and touching bids are carried over at the IMM. In the Hertz example, then, Citigroup, Credit Suisse, Deutsche Bank, and RBS Capital Markets all carry over a limit buy order for $2 million at 25.375.

used this final price as the recovery amount, that is, protection buyers received from protection sellers the notional amount of their CDS × (100% − 26.375%), or 73.625%. Counterparties who chose physical settlement worked through their dealers as follows. Protection buyers delivered deliverable obligations of their choice for 26.375 and received a cash settlement of 73.625, for total proceeds of 100. Protection sellers paid 26.375 for any bond delivered to them and paid a cash settlement of 73.625, again, for total proceeds of 100.

This section concludes by noting that, while CDS auctions have generally been successful in settling at reasonable final prices, success is not assured. A recent example of failure was the auction to settle Europcar Mobility Group CDS on January 13, 2021. Just before the auction, deliverable bonds were trading at about 70, and phase one of the auction resulted in an IMM of 73 and a NOI to buy €43.3 million. But in phase two, limit orders to sell totaled only €35.9 million. Hence, with no price high enough to match limit sell orders to the NOI to buy, the final auction price, by rule, was set to 100. Buyers of CDS protection received nothing. Some commentators attributed the lack of limit sell orders to the fact that Europcar restructuring agreements prevented many bondholders from selling their bonds through the CDS auction. Other commentators, however, claimed that there were more than enough bonds available to cover the NOI, but potential sellers simply did not show up at the auction.[23]

14.12 OPPORTUNISTIC CDS STRATEGIES

In theory, CDS traders take the behavior of debt issuers as given, and debt issuers essentially ignore CDS trading on their obligations. In practice, however, because of the complexities of CDS contracts (e.g., specifying events of default; defining reference entities and deliverable obligations; setting rules for settlement auctions), both CDS traders and issuers have occasionally tried to profit from crossing the lines separating their respective markets. These instances have been called *opportunistic CDS strategies*.

A particularly well-known example is the case of the real estate company, Hovnanian, and its "manufactured default." In February 2018, Hovnanian and an asset manager that had bought over $300 million of protection on Hovnanian conducted the following transactions:

1. The asset manager loaned Hovnanian a significant sum at below-market rates.
2. Hovnanian conducted a bond exchange offer, in which it issued new, very long-term debt that would sell at a significant discount, given its relatively low credit rating; and in which some of its higher-priced outstanding bonds were purchased by one of its own subsidiaries.

[23] Levine, M. (2021), "Europcar," *Money Stuff*, Bloomberg, January 15.

3. Hovnanian committed to skip the coupon payment on its bonds owned by that subsidiary. This default would not trigger cross-default provisions, but would trigger a credit event with respect to Hovnanian CDS.[24]

Through these transactions, Hovnanian was able to borrow at advantageous rates, and the asset manager, if the credit event occurred, would benefit not only from the triggering of its long CDS position, but also from the issuance of the low-priced bond issue, which, because it was deliverable into the CDS, would lower the CDS settlement auction price and increase the payout to protection buyers. While Hovnanian did skip the coupon payment per its commitment, the company ultimately reversed course and paid before the end of the 30-day grace period. Pressure from regulators and from asset managers who had sold protection on Hovnanian, along with a settlement between the asset managers on both sides of the CDS, led to the abandonment of this manufactured default.

The Hovnanian case threatened confidence in the integrity of the definition of "default" in the CDS market. In response, ISDA changed the definition of "failure to pay" to include the language, "it shall not constitute a Failure to Pay if such failure does not directly or indirectly either result from, or result in, a deterioration in the creditworthiness of financial condition of the Reference Entity."[25]

While this change in definition is likely to eliminate some opportunistic CDS strategies, other possibilities remain. With respect to the buyers of CDS protection influencing default, consider the case of the telecommunications company, Windstream. In 2015, in the course of a restructuring, the company seems to have violated an existing bond covenant. No one objected until 2017, however, when a hedge fund bought some of the bonds and sued to put the company into default. Many market participants believe that the hedge fund bought CDS protection on Windstream, bought the bonds, and then sued to collect on the CDS. And the hedge fund did prevail in court, despite other bondholders' willingness to waive the alleged covenant violation. In this case then, not only did the event of default not reflect underlying credit conditions, but CDS buyers also seem to have precipitated that default.[26]

---

[24]Cross-default provisions say that a default on one issue triggers a default on many or all of an issuer's other obligations. These provisions protect the holders of one bond issue, who happen to have not yet experienced a default, from being disadvantaged by the cure of another issue's default. A simple example is the case of one issue defaulting before another, simply because of a relatively early payment due date.

[25]ISDA (2019), "2019 Narrowly Tailored Credit Event Supplement to the 2014 ISDA Credit Derivatives Definition," July 15, p. 3.

[26]For one account of these events, see Levine, M. (2019), "Windstream," *Money Stuff*, Bloomberg, February 19.

Opportunistic strategies seem also to have been pursued by sellers of CDS protection. In 2018, an asset manager, widely believed to have sold protection on the publishing company, McClatchy, made the following agreement with the company. The asset manager would buy newly issued bonds from a subsidiary of McClatchy, with the proceeds used to refinance (and extinguish) the existing bonds of McClatchy itself. Because outstanding CDS included McClatchy, but not its subsidiary, as reference entities, McClatchy CDS would be "orphaned" – without any underlying debt, McClatchy could not possibly default. Furthermore, protection buyers would still have to make premium payments. News of this agreement dramatically reduced the value of protection, generating profits for sellers of protection. As it turned out, the proposed deal did not go through and, in early 2020, McClatchy filed for bankruptcy. And, in the subsequent CDS auction, the final price was two; that is, McClatchy CDS eventually paid 98 cents on the dollar.[27]

As a final example of opportunistic CDS strategies, in this instance with the issuer as first mover, consider the case of Sears Holding Corp. At the time of its bankruptcy, in late 2018, there was a lot of CDS outstanding not on the holding company, but on its subsidiary, Sears Roebuck Acceptance Corp (SRAC). But relative to SRAC CDS, there was relatively little SRAC debt outstanding, perhaps about $200 million. Therefore, CDS buyers were willing to pay a premium to buy SRAC debt: they could thus ensure having enough bonds to sell into the settlement auction, which, in turn, would keep the final auction price low and default compensation high. Recall that in the Europcar auction, with too few submitted limit sell orders, the final auction price was 100 and protection buyers received nothing. In any case, Sears realized that – because of the CDS market – the market value of SRAC bonds far exceeded the expected discounted value of their cash flows and offered to sell a previously overlooked $900 million of intercompany SRAC notes. CDS buyers were outbid in this sale, however, by CDS sellers – who wanted to prevent these SRAC bonds from being delivered into the settlement auction. In the end, however, to stop the back-and-forth litigation, CDS buyers and sellers came to a settlement on their own.

There are several examples of opportunistic CDS strategies other than the ones described here, but at least two lessons emerge. First, market participants need to be aware of special situations that could cause final recovery to differ from expectations. Second, if opportunistic strategies were ever to dominate trading, many market participants might very well abandon the market, believing that it no longer offered straightforward transfers of credit risk.

[27]For more detail on these events, see Carruzzo, F., and Zide, S. (2018), "Opportunistic CDS Strategies Available to CDS Protection Sellers Part II: McClatchy and Sears," *Debt Dialogue*, Karmer Levin, June 6.

14.13 CASE STUDY: THE LONDON WHALE

In the first half of 2012, JPMorgan Chase (JPM) lost more than $6 billion on its "London Whale" trades. With assets at the time of over $2 trillion and 2012 income of $21 billion, the loss in no sense endangered the bank, but it was large and embarrassing. In this chapter, the case is used to explain more fully the workings of index CDS and CDS *tranches*, though other lessons include the need to define the purpose and scope of hedging programs and the importance of sizing trades so as not to "become the market."[28]

Index CDS and Tranches

The important instruments for understanding the events of the case are CDX.NA.IG.9 and its tranches. As described earlier in the chapter, CDX.NA.IG is an equally weighted index of 125 investment-grade North American entities. The ".9" indicates a particular series of this index, which came out in the third quarter of 2007. Hence, 10-year protection on this index expired in December 2017 and five-year protection in December 2012. As the London Whale events unfolded in the first quarter of 2012, 10-year CDS on this index matured in just under six years, and five-year CDS on the index matured in just under a year.

As of April 2012, four of the 125 names in CDS.NA.IG.9 had experienced credit events during the financial crisis of 2007–2009 and, therefore, had dropped out of the index: CIT, a subprime mortgage lender; WAMU, a Savings and Loan Association; Fannie Mae, a government sponsored mortgage agency; and Freddie Mac, a government sponsored mortgage agency. Table 14.17 shows the impact of these four credit events on the index. The second column shows the final price in the CDS auction for each of the names. To understand the third column, consider $125 million

[28]The presentation here relies on the following sources: Childs, M., Ruhle, S., and Harrington, S. (2012), "Blue Mountain Said to Help Unwind JPMorgan's Whale Trades," *Bloomberg Business*, June 21; Tyler, D. (2012), "Behind 'the Iksil Trade' – IG9 Tranches Explained," *Zero Hedge*, April 10; JPMorgan Chase (2013), "Report of JPMorgan Chase & Co. Management Task Force Regarding 2012 CIO Losses," January 16; Pollack, L. (2012), "Recap and Tranche Primer," *FT Alphaville*, May 16; Pollack, L. (2012), "The High Yield Tranche Piece," *FT Alphaville*, May 17; Pollack, L. (2012), "Unwind? What Unwind? – Part 2," *FT Alphaville*, June 27; Protess, B., Sorkin, R., Scott, M., and Popper, N. (2012), "In JPMorgan Chase Trading Bet, Its Confidence Yields to Losses," *The New York Times*, May 11; Ruhle, S., Keoun, B., and Childs, M. (2012), "JPMorgan Trader's Positions Said to Distort Credit Indexes," *Bloomberg Business*, April 6; US Senate (2013), "JPMorgan Chase Whale Trades" and "Exhibits," March 15; and Zuckerman, G., and Burne, K. (2012), "London Whale Rattles Debt Market," *The Wall Street Journal*, April 6.

TABLE 14.17 Names Dropped from CDX.NA.IG.9, as of April 2012. Index Loss in Percent.

Name	Price in CDS Auction	Index Loss
CIT	68.125	0.255
WAMU	57.000	0.344
Fannie Mae	91.510	0.068
Freddie Mac	94.000	0.048
Total		0.715

original notional amount of the index, which would include \$1 million notional amount in each of the underlying 125 names. Given an auction price of 68.125 for CIT obligations, \$1 million notional amount of CDS on CIT lost (100% – 68.125%) × \$1 million, or \$318,750, which is 0.255% on the index notional amount of \$125 million. Similar calculations for the rest of the names give the rest of the third column, which sum to a total index loss of 0.715%, or \$893,750 on the original \$125 million index notional amount. In summary then, over the life of \$125 million notional amount of this index, protection sellers paid protection buyers \$893,750 for four credit events, and the remaining notional amount fell to \$121 million. Additionally, of the remaining names in the index, five had CDS spreads of over 500 basis points and four has spreads of about 400 basis points. In other words, while issued as an investment-grade index, this series experienced significant credit deterioration.

As for CDS index tranches in the context of CDX.NA.IG.9, Table 14.18 describes compensation payments and Table 14.19 describes premium payments and pricing. Starting with Table 14.18, note that the index notional

TABLE 14.18 Structure of Tranches on CDX.NA.IG.9, as of April 2012. Attachment and Detachment Points Are in Percent. Notional Amounts Are in Dollars.

	Before Losses			After Losses		
Tranche	Attach	Detach	Notional	Attach	Detach	Notional
Equity	0	3	3,750,000	0	2.361	2,856,250
Jr. Mezz.	3	7	5,000,000	2.361	6.493	5,000,000
Sr. Mezz.	7	10	3,750,000	6.493	9.592	3,750,000
Sr.	10	15	6,250,000	9.592	14.757	6,250,000
1st Super Sr.	15	30	18,750,000	14.757	30.253	18,750,000
2nd Super Sr.	30	100	87,500,000	30.253	100.00	84,393,750
Index			125,000,000			121,000,000

in this example is fixed at $125 million, and consider the Senior Mezzanine tranche. Its original, "Before Losses" *attachment* point of 7% means that sellers of protection on this tranche do not have to make any compensation payments unless losses on the index exceed 7%, or $8.75 million on the $125 million index notional. If losses are 8%, for example, sellers of protection on the Senior Mezzanine tranche have to pay protection buyers on this tranche (8% − 7%) × $125,000,000, or $1.25 million. The tranche's "Before Losses" *detachment* point of 10% means that sellers of protection stop making compensation payments when losses exceed 10% or $12.5 million. In fact, the original notional amount of the Senior Mezzanine tranche is $3.75 million, precisely because sellers of protection on this tranche have no further responsibilities after paying protection buyers (10% − 7%) × $125,000,000, or $3.75 million.

Now consider the equity tranche. From the "Before Losses" columns, protection sellers on this tranche have to pay protection buyers from the first loss – the attachment point of 0% – until losses reach 3%. The "After Losses" column shows what happened to the equity tranche given the losses to CDX.NA.IG.9 through April 2012, as described in Table 14.17. With the index losing 0.715%, sellers of protection on the equity tranche had to pay protection buyers 0.715% of $125 million, or $893,750. Hence, the notional amount of the equity tranche was reduced from its original $3,750,000 to the 'After Losses" notional of $3,750,000 minus $893,750, or $2,856,250. Furthermore, with these losses arising from credit events in four names, each with $1 million notional amount, the notional of the index referenced by the tranches fell from the original $125 million to $121 million. Finally, because sellers of protection on the equity tranche can be obliged to pay no more than their remaining notional amount of $2,856,250, the "After Losses" detachment fell from the original 3% to $2,856,250/$121,000,000, or 2.361%.

With protection sellers of the equity tranche paying compensation as just described, the attachment and detachment points of the rest of the tranches shifted as well. Protection sellers on the Junior Mezzanine tranche start paying where the equity tranche leaves off, at 2.361%, and stop at 6.493%, when their notional of $5 million is exhausted: $(6.493\% - 2.361\%) \times \$121{,}000{,}000 = \$5{,}000{,}000$. The new attachment and detachment points of the remaining tranches are determined similarly. As for the "After Losses" notional amounts, that of the equity tranche falls, as explained, while all others, which have not yet experienced losses, remain the same, except for the most senior tranche. With only $121 million notional amount of the index remaining, and with the notional amounts of all the other tranches set as just described, the absolute worst-case loss for the Second Super Senior tranche is a loss of the residual notional amount of $84,393,750. (While the attachment point of this tranche increases from 30% to 30.253%, it starts to bear losses at the same dollar amount as before: $30\% \times \$125{,}000{,}000 = \$37{,}500{,}000$ is the same as \$893,750 – the losses already incurred – plus the new dollar attachment point of $30.253\% \times \$121{,}000{,}000 = \$36{,}606{,}250$.)

Table 14.18 includes all tranches so as to compare their cash flows given credit events, but tranches exist independently of each other. A particular tranche with a particular notional amount is created when a protection seller trades that tranche with a protection buyer. CDS index tranches are *synthetic* in the sense that there is no underlying portfolio of actual index CDS. The underlying CDS index is used solely to determine the compensation payments of the tranches.

Moving to premium payments and pricing, Table 14.19 repeats the tranche names and original attachment points and gives the coupon and upfront amounts for each as of mid-May 2021. Not surprisingly, protection buyers pay more, and protection sellers receive more, for the more junior tranches, which experience losses earlier. A buyer of protection on \$1 million notional of the 3–7% or Junior Mezzanine tranche, for example, pays a running 500 basis points or \$50,000 per year plus an upfront amount of 29.43% or \$294,300. By contrast, a buyer of protection on \$1 million notional of the 15–30% or First Super Senior tranche pays a running \$10,000 plus an upfront amount of \$6,700.

Table 14.19 also reports the delta of each tranche, expressed in index equivalents. For example, the sensitivity of the 7–10% tranche to changes in credit spreads is 5.18 times the sensitivity of the underlying index. The 15–30% tranche, on the other hand, is 1.30 times as sensitive as the index. The delta of the tranches tend to decline with seniority, from the 3–7% to the 30–100% tranche: the higher the seniority, the less likely losses are to impact the tranche, and the lower the delta. The delta of the equity tranche, however, is less than that of the 3–7% and 7–10% tranches. Imagine for a moment that the remaining equity tranche notional was almost certain to be wiped out by future losses. In that case, marginal increases or decreases in credit spreads would not change the price of the equity tranche by much.

TABLE 14.19 Pricing of Tranches on CDX.NA.IG.9 Maturing in December 2017, as of Mid-May 2012. Attachment, Detachment, and Upfront Amounts Are in Percent. Coupon Is in Basis Points.

Tranche	Attach	Detach	Upfront	Coupon	Delta
Equity	0	3	72.71	500	3.92
Jr. Mezzanine	3	7	29.43	500	6.58
Sr. Mezzanine	7	10	6.85	500	5.18
Sr.	10	15	9.80	100	2.96
1st Super Sr.	15	30	0.67	100	1.30
2nd Super Sr.	30	100	−3.88	100	0.29

Those same marginal changes to credit spreads would, by contrast, have a significant impact on the as yet untouched 3–7% tranche. The situation in Table 14.19, although not nearly as extreme, is analogous: the high coupon and the high upfront payment of the equity tranche indicate a very high expectation of large losses in the future. Part of this, by the way, is due to the nine high-spread names in the index at the time, as mentioned earlier.

While not directly evident from the table, the market's views on the correlations of defaults across names in the index is a significant determinant of relative tranche value. Given the probabilities of default for each name, higher correlations of those defaults imply more defaults and a higher likelihood of wiping out the more junior tranches and of inflicting some losses on more senior tranches. Moving from low to moderate correlations, therefore, would impact the prices of the junior tranches most. Moving from moderate to high correlations, however, might have a particularly large impact on the prices of the senior tranches, which had before seemed untouchable.

The Synthetic Credit Portfolio (SCP)

With this introduction to CDS indexes and tranches complete, the text turns to the case proper, beginning with the chief investment office (CIO) of JPM as of the end of 2011. The job of the CIO was to invest funds in a wide range of high-quality fixed income products, with the objectives of meeting the future liquidity needs of the bank and of earning a reasonable rate of return. The amount of funds available to the CIO had increased dramatically, from $76.5 billion in 2007 to $365 billion in 2011, with the greatest increase during the financial crisis of 2007–2009 as JPM, with its "fortress balance sheet," attracted about $100 billion of deposits.

The CIO launched its Synthetic Credit Portfolio (SCP) in 2007 to protect the bank's investments and loans from adverse credit scenarios. The SCP was often short credit (i.e., taking positions that do well when credit conditions deteriorate), but sometimes sold CDS protection and was sometimes overall long credit. In hindsight, the CIO can be criticized for not identifying a particular portfolio of assets it was hedging with SCP and for not articulating a particular hedging policy. In any case, SCP did earn profits: somewhere between $200 million and $625 million in 2007; $170 million in 2008; $1.05 billion in 2009, primarily from a bet on General Motors' bankruptcy; $149 million in 2010, while consciously shrinking the book; and $453 million in 2011, mostly due to having bought protection on American Airlines that expired just three weeks after its bankruptcy.

With this history, the SCP ended 2011 with $51 billion of net CDS notional. It was overall short credit, though it was both long and short particular indexes and tranches, and it had on the relative value trade of being long investment-grade credits and short high-yield credits. Various documents and correspondence made clear that SCP intended to trade credit risk actively, as opposed to hedging exposures held elsewhere in the bank. It had developed a bullish view on the economy and planned, in 2012, to establish a long credit position. It also planned to continue to trade opportunistically, for example: to maintain "smart" shorts on credit; it "likes cheap options"; profit from defaults without paying "too much"; and repeat the successful trade on American Airlines but avoid losses like the $50 million lost on Kodak's default in January 2012. Finally, going into 2012, the CIO realized that its positions were becoming so large that unwinds would incur extremely large transaction costs.[29]

The four active strategies pursued by SCP in 2012 all performed poorly:

1. SCP went long credit, changing its sensitivity to credit spreads from a relatively mild short in late 2011 to a long of $60 million per basis point. This positioning did well as credit spreads fell through the end of March, but spreads then rose dramatically, which inflicted significant losses. SCP expressed its long position mostly by buying CDX.NA.IG.9. Because this index contained several high-yield names, as mentioned before, its purchase would also hedge some of SCP's existing high-yield shorts. While in some ways logical, this was a dangerous choice. First, as mentioned previously, SCP positions were already becoming too large relative to the liquidity of the market. Therefore, buying CDX.NA.IG.9 to hedge shorts, instead of just buying back those shorts, made the portfolio even larger and more unwieldy. Second, SCP's oversized position in CDX.NA.IG.9 exposed it to predatory trading by other market participants. As a result, as discussed further presently, CDX.NA.IG.9 underperformed its constituent CDS so that SCP lost more money from the increase in credit spreads than it might have otherwise.
2. SCP continued its relative value trade of being long investment-grade risk and short high-yield risk. This strategy lost money as the difference between high-yield and investment-grade yields fell from 560 basis points in December 2011 to 480 basis points in March 2012, though the difference recovered to 527 basis points by mid-April.
3. SCP put on an investment-grade flattener, betting that the difference between long-term investment-grade spreads and short-term investment-grade spreads would narrow. SCP chose to implement this strategy with IG.9 as well, using the once 10-year and once five-year expirations mentioned already. This strategy lost money too, with the spread between the 10- and five-year CDS indexes rising from 16 basis points in December 2011 to 48 basis points by mid-April.

[29]SCP attributed a $15 million loss in January 2012 to trying to unwind high-yield shorts. It was estimated that reducing risk-weighted assets by $25 billion would cost between $400 million and $590 million and that reducing positions by 25–35%, pro rata, would cost of more than $500 million, including lost revenues.

4. SCP sold the "belly" of CDS index tranches and bought the equity tranche and the index itself. Again SCP favored IG.9 tranches, because they had historically traded with relatively high liquidity. SCP lost money here as well. Focusing on IG.9, SCP was short mostly in the 10–15% and the 15–30% tranches, while long the index and the 30–100% tranche. As credit spreads rose, both the index and all the tranches fell in value. But the relative price changes were surprising. The price of the 30–100% tranche fell a lot, as if the names in the index were highly correlated, while the 10–15% and 15–30% tranches fell less, as if there were limited increases in correlation. SCP lost money, therefore, because it was long the tranche that underperformed and short the tranches that outperformed. As discussed presently, this price action may very well have been due to SCP having an oversized position in IG.9 and its tranches.

Losses across these strategies were $718 million in the first quarter of 2012 and an additional $1.5 billion loss in April. Management of the SCP portfolio was taken away from the CIO at the end of April, but the positions continued to lose money, mounting to a cumulative loss through the end of June 2012 of $5.8 billion. At this point, many of the positions were moved to other parts of the bank, which made it impossible for the public to determine the ultimate losses from the portfolio.

The raw size of SCP positions magnified its losses. SCP's total CDS notional amount rose from $51 billion at the end of 2011 to $157.1 at the end of March 2012. These positions were also very large relative to trading volume. Its position in IG.9 indexes at that time was over 10 days of average daily trading volume, and its position in one of the iTraax series was over eight times average daily volume. The impact of these large positions can be seen both from market prices and from the behavior of market participants at that time.

With respect to market prices, consider the 10-year IG.9 *skew*, which is defined as the difference between the spread of the underlying CDS and the 10-year IG.9 CDS spread. The more positive the spread, the richer the IG.9 spread relative to its constituents. This skew rose from about 10 basis points in July 2011, reached a high of 35 basis points in fall 2011, and was about 25 basis points at the start of 2012. In hindsight, the richness of the IG.9 index skew was likely caused by SCP's aggressive buying of the index.

With respect to the behavior of market participants, hedge funds had noticed the skew and, into 2012, bought the cheap constituents and sold the rich IG.9. They knew, however, that SCP's large index purchases might very well keep the skew high, which turned out to be the case. News articles appeared in early April, likely encouraged by these funds, mentioning the large positions of JPM's CIO, attributing the persistent skew to these positions, and questioning whether the incoming Volcker rule would force

JPM to liquidate these positions. These articles, together with some efforts by SCP to reduce positions, encouraged market participants to bet on an upcoming liquidation, that is, to sell IG.9, which SCP would ultimately have to sell, and to buy the constituent CDS as a hedge. This activity, narrowed the skew, of course, causing IG.9 to underperform and to inflict additional losses on SCP's portfolio. At first, JPM responded to the news articles by saying that the CIO's positions could be held indefinitely. But with credit spreads rising, SCP losses mounting, and the market increasingly aware of SCP's predicament, the hedge funds won. SCP's portfolio would be at least partially unwound, and the skew fell to zero by July 2012. In fact, some hedge funds unwound their IG.9 shorts – at a profit – by buying IG.9 directly from SCP as part of its unwind.

CHAPTER 15

Mortgages and Mortgage-Backed Securities

A mortgage is a loan collateralized by property. This chapter focuses on residential mortgages, through which prospective homeowners borrow money to purchase a home on the collateral of that home.

The most common mortgage loans in the United States are for 30 years and carry fixed rates of interest. They require equal monthly payments, with each allocated in part to interest (which in many cases is tax deductible) and in part to paying down principal. Mortgages also give homeowners the option to pay off or *prepay* all or part of the outstanding balance at any time. An implication of these contractual features is that, when mortgage rates fall, homeowners can profitably *refinance* their existing, high-rate mortgages by prepaying them with money raised from new, lower-rate mortgages. Homeowners also have to prepay when they sell a property, because the property is the collateral behind the mortgage. Details are discussed throughout the chapter, but the extent to which homeowners prepay is a key risk factor for lenders and investors: mortgages that carry an above-market rate lose value when homeowners prepay, while those that carry a below-market rate gain value with prepayments.

Another key risk factor is default. While mortgages are collateralized by homes, lenders can lose principal if a homeowner defaults when the value of the home is less than the outstanding balance on the mortgage. Lenders protect themselves against this scenario by lending less than the value of the home at the time of purchase, that is, by requiring that the original *loan-to-value* (LTV) ratio be less than some amount, typically 80%. The resulting buffer may not turn out to be sufficient, however, if the value of the home falls dramatically. Furthermore, with respect to any residual mortgage liability above the value of the home, lenders may or may not have recourse to the homeowner's other assets, depending on the laws of the relevant state.

The first section of the chapter describes the mortgage market in the United States, and the large role played by government agencies. Sections 15.2 to 15.4 introduce the payment conventions of mortgage loans, both fixed and floating rate, and the value of the prepayment option for fixed-rate loans. Sections 15.5 to 15.7 discuss mortgage pools, the most basic *mortgage-backed security* (MBS), along with their pricing and risk profiles. Sections 15.8 to 15.10 present the *TBA* and *specified pools* markets, and how they are used for hedging and financing. Section 15.11 deals with other types of MBS and Section 15.12 explains *credit risk transfer* (CRT) securities.

15.1 THE MORTGAGE MARKET IN THE UNITED STATES

Historically, banks dominated mortgage lending. They had existing relationships with potential borrowers (i.e., depositors, business owners) and their deposit base provided a relatively stable source of funds for making and holding mortgage loans. Since the Great Depression, however, US government policy has cultivated a secondary market in mortgages. The 1930s saw the creation of the Federal Housing Administration (FHA) and the Federal National Mortgage Administration, or Fannie Mae (FNMA). The FHA was created to insure lenders – for a fee – against the loss of principal or interest on approved mortgages, and FNMA was created to trade and hold FHA-insured loans. The Veterans Administration (VA) was created in the 1940s to offer veterans federally insured mortgages, which were added to FNMA's remit. The Government National Mortgage Association – Ginnie Mae (GNMA) – and the Federal Home Loan Mortgage Corporation – Freddie Mac (FHLMC) – were created in 1968 and 1970, respectively. GNMA's mission is to package FHA-, VA-, and other government-insured mortgages into MBS; to collect fees to insure those MBS; and to sell those MBS to investors. FHLMC's original purpose was to create, insure, and sell MBS from mortgages purchased from savings and loan associations.[1]

In their simplest form, MBS are large portfolios or *pools* of individual mortgages that *pass through* payments of interest and principal from borrowers to investors. There were several ideas behind the *securitization* of mortgages through MBS. First, enabling lenders to sell their mortgage loans facilitates the movement of investment funds from where they are to where they are needed. Second, a large and diversified pool of mortgages trades with much greater liquidity than would individual mortgages. Third, the guarantees described in the previous paragraph significantly increase the

[1]For a concise history, see, for example, Federal Housing Finance Agency (2011), "A Brief History of the Housing Government-Sponsored Enterprises," in Inaugural Semiannual Report to the Congress, March 31, pp. 47–54.

attractiveness of mortgages, because investors do not have to research the creditworthiness of individual homeowners nor bear that credit risk. Fourth, government can use its role in insurance and securitization to direct mortgage capital according to various policy goals, like affordable housing.

In the decades preceding the financial crisis of 2007–2009, FNMA and FHLMC were transformed into *government-sponsored enterprises* (GSEs): they became privately owned, publicly traded corporations, but were heavily regulated, received many benefits from their quasi-government status, and were required to pursue various public policy objectives. They had two business lines: the guarantee business, which made money by charging fees to package and insure MBS; and the portfolio business, which made money by borrowing in public debt markets at relatively low rates – available because of their quasi-government status – and investing in mortgages and MBS.

GNMA, FHLMC, and FNMA are collectively called "the agencies," and their MBS are all called "agency MBS," even though GNMA is the only one of the three that is actually a government agency and explicitly backed by the "full faith and credit" of the federal government. In any case, from about 2000 to the financial crisis of 2007–2009, private banks, investment banks, and other mortgage lenders dramatically increased their issuance of *non-agency* or *private-label* MBS, particularly to pool together mortgages that were larger or less creditworthy than those acceptable for agency MBS. In fact, just before the crisis, non-agency MBS issuance exceeded agency issuance.

During the financial crisis of 2007–2009, housing prices collapsed and many homeowners defaulted on their mortgages. Private-label MBS, which held particularly low-quality mortgages, suffered devastating losses, but the GSEs suffered as well. With losses on their guarantee and portfolio businesses threatening their viability, they were rescued by the US Treasury and put into *conservatorship*, in which state they remain to this day. Before the crisis, claims on the GSEs were not legally backed by government, but it was widely believed – and market prices reflected – that, in a crisis, the government would back those claims. And that belief was fulfilled: no holder of GSE debt or MBS lost principal or interest in the crisis.[2] In this sense, the situation today is as it was before the crisis: even in conservatorship, GSE claims are not explicitly backed by the government but, should the need arise, markets expect these claims to be honored by the government.

Under conservatorship, the GSEs have continued to grow in overall size, but have been subject to a number of policy changes. First, their portfolio businesses are limited in size and scope to the facilitation of their guarantee

[2]The establishment of the conservatorship itself did trigger a credit default swap auction on GSE debt, and final auction prices were less than par. Nevertheless, bonds issued by the GSE made timely payments of all interest and principal.

TABLE 15.1 First-Lien Mortgages, Gross Issuance, in $Trillions.

	2020		Q1–Q3 2021	
	Issuance	**%**	**Issuance**	**%**
GSE MBS	2.390	59.2	2.06	55.5
GNMA MBS	0.742	18.4	0.57	15.4
Bank Portfolio	0.869	21.5	1.02	27.5
Private-Label MBS	0.038	0.9	0.06	1.6

Source: Ginnie Mae (2021), "Global Markets Analysis Report," various months.

businesses. Second, they offload some of their credit risk to the private sector through the sale of *credit risk transfer* securities, which are the subject of the last section of this chapter. Third, to increase liquidity, they are directed to keep their MBS similar enough to be fungible. In particular, since June 2019, MBS forward contracts, namely, *TBAs*, are written on *uniform MBS* (UMBS), which designation allows for the delivery of both FHLMC and FNMA MBS. Fourth, and most recently, provisions have been made for the GSEs to build capital in preparation for an exit from conservatorship. There are no plans in place, however, for actually effecting such an exit.

This brief history sets the stage for describing the US mortgage market as it stands today. Tables 15.1 and 15.2 show the extent to which agency MBS dominate the US residential mortgage market. Table 15.1 reports the gross issuance of first-lien mortgages in 2020 and in the first three quarters of 2021.[3] Starting with 2021, 70.9% of new mortgages are sold through agency

TABLE 15.2 Residential Mortgages Outstanding (1–4 Family), as of September 2021, in $Trillions.

	Outstanding	**%**		**%**
Agency MBS	8.18	66.7	FNMA	40
			FHLMC	34
			GNMA	26
Unsecuritized 1st Lien	3.30	26.9		
Private-Label MBS	0.39	3.2		
Home Equity Loans (2nd Lien)	0.40	3.3		
Total	12.27	100.0		100

Sources: Federal Reserve Board, Financial Accounts of the United States; and Ginnie Mae (2021), "Global Markets Analysis Report," December 2021.

[3] "Gross" issuance includes refinancings. But because refinancings extinguish existing mortgages, they do not necessarily add to net outstanding mortgage principal.

MBS; 27.5% are made and held by banks; and only 1.6% are sold through private-label MBS. The dominance of agency MBS was even greater in 2020, perhaps due to the pandemic and economic shutdowns, during which time private credit conditions were strained. Table 15.2 shows the prevalence of agency MBS with respect to amounts outstanding. About 67% of mortgages are held in agency MBS, and only 3.2% in private-label MBS. The rightmost columns of the table separate outstanding agency MBS volumes by agency.

Because the agencies play such a large role in the mortgage market, their underwriting criteria play an outsized role as well. Government-insured mortgages, like those made by the FHA and VA, are called *government* loans. All other loans are called *conventional* loans. Mortgages eligible to be included in agency MBS are called *conforming* mortgages, while those not eligible are called *nonconforming*. For the GSEs, for example, at the time of this writing, conforming mortgages must have a FICO score greater or equal to 620; an LTV less than 97%; a debt-to-income ratio (DTI) no greater than 45%; and, for most geographical areas, a loan amount less than \$647,200.[4] The requisite characteristics for FHA and VA mortgages are different. In any case, *jumbo* mortgages, which have loan amounts greater than the "conforming balance limit," along with mortgages of lower than conforming credit quality, have to be funded outside the agency ecosystem.[5] In addition, to the extent that the GSEs limit purchases of mortgages with certain characteristics, like those funding second homes or investment properties, private markets have the opportunity to satisfy the residual demand.

Table 15.3 shows averages of selected characteristics of loans to first-time homebuyers for the purposes of purchasing a home (rather than refinancing an existing mortgage). Loan characteristics do not vary much across FNMA and FHLMC, because, as mentioned before, their business models are converging. GNMA loans are smaller and, by standard metrics, less creditworthy, with significant variation across FHA, VA, and other government-guaranteed loans.

Deposit franchises give banks competitive advantages in making and holding mortgage loans. These advantages are much less pronounced, however, for *originating* conforming mortgages, that is, for arranging mortgages

---

In a first-lien mortgage, the lender has first or senior rights to the property. A homeowner's primary mortgage is a first lien, while any additional loans taken on the same property, like loans to make the required down payment and home equity loans, are second liens and junior to the first lien.

[4]FICO is a credit metric created by the Fair Isaac Corporation that increases with measured borrower creditworthiness, to a maximum of 850. DTI is defined as monthly debt payments divided by monthly income.

[5]Jumbo loans were between 3% and 4% of all loans from 2018 through the first half of 2020. See Federal Housing Finance Agency (2021), "What Types of Mortgages Do Fannie Mae and Freddie Mac Acquire?" April 14, pp. 2–3.

TABLE 15.3 Average Characteristics of First-Time Homebuyer Purchase Loans, as of December 2020.

				Within GNMA MBS		
	FNMA	**FHLMC**	**GNMA**	**FHA**	**VA**	**Other**
Loan Amount ($)	305,134	264,303	245,013	237,630	299,448	174,011
FICO Score	750.4	744.7	688.5	678.6	713.3	701.1
LTV (%)	87.7	88.0	96.8	95.5	99.6	99.4
DTI (%)	34.1	34.4	41.5	43.1	39.5	34.9

Source: Ginnie Mae (2021), "Global Markets Analysis Report," February.

with borrowers and then selling those mortgages into agency MBS. With the increasing dominance of agency MBS, therefore, the market share of non-bank originators has steadily increased, from less than 25% in 2007 to more than 75% toward the end of 2021.[6] In fact, a rough calculation assuming a 25% bank share of agency MBS originations, and mortgage gross issuance as described in Table 15.1, that is, 71% for agency MBS and 27.5% for bank loan portfolios, gives an overall bank participation in mortgage issuance of 25% times 71% plus 27.5%, or about 45%.

Nonbank origination *channels* are classified into three groups: retail, broker, and correspondent. Retail originators contact borrowers directly. Broker originators provide various services to independent brokers who seek out and find potential borrowers. And correspondent originators buy loans from smaller banks and credit unions.

This market overview concludes with *mortgage servicers*. Servicers receive ongoing fees to administer mortgage loans; keep in touch with borrowers; pass payments from borrowers to investors; help borrowers through loan modifications or forbearance; and, in the case of missed payments, temporarily advance funds to investors. Lenders can either keep the *servicing rights* to a loan or sell the rights to another entity. Arguments for keeping the rights include their performing as a hedge to an origination business. First, servicing provides steady income through periods of slow origination. Second, when refinancings decrease servicing income (by extinguishing existing mortgages), income from refinancing originations increases. Another reason originators keep servicing rights is to keep in touch with borrowers who may be contemplating refinancing. In that way, lenders can *recapture* refinancing, that is, market their own refinancing products with the goal of preventing borrowers from refinancing with a different originator.

[6]See Fritzdixon, K. (2019), "Bank and Nonbank Lending over the Past 70 Years," *FDIC Quarterly* 13(4), pp. 31–39; and Ginnie Mae (2021), "Global Markets Analysis Report," December, p. 26.

15.2 FIXED-RATE MORTGAGE LOANS

Though terms of 10, 15, and 20 years are also available, the most common mortgage is a 30-year, fixed-rate, *level payment* mortgage. For example, a homeowner might borrow \$100,000 from a bank at 4.5% and agree to make payments of \$506.6853 every month for 30 years. (The extra decimal places are given to help the reader reproduce the calculations to follow.) The loan rate and the monthly payment are related by,

$$\$506.6853 \sum_{n=1}^{360} \frac{1}{\left(1+\frac{4.50\%}{12}\right)^n} = \$100{,}000 \tag{15.1}$$

$$\$506.6853 \frac{12}{4.50\%}\left[1 - \frac{1}{\left(1+\frac{4.50\%}{12}\right)^{360}}\right] = \$100{,}000 \tag{15.2}$$

where Equation (15.2) follows by applying Equation (A2.16).

These equations say that, at a rate of 4.5%, the present value of the monthly payments on the mortgage over its 30-year or 360-month life equals the initial loan amount of \$100,000. This relationship is simply a market convention to define a monthly payment from a mortgage rate, just as price is quoted given yield to maturity or *vice versa*. Truly pricing a mortgage loan is much more complicated, of course, with the need to account for a term structure of risk-free rates, the value of the prepayment option, and a spread or term structure of spreads to reflect credit risk.

The market convention of a single rate to describe a mortgage loan is also used to divide its monthly payments into interest and principal components. Table 15.4 gives selected rows of the *amortization table* for the mortgage just described, with the monthly payment rounded to \$506.69. The starting balance is \$100,000. The interest component of the first payment is the interest on that starting balance at a rate of 4.50%, that is, $\$100{,}000 \times 4.50\%/12 = \$375.00$. The principal component of the first payment is simply the total monthly payment, fixed at \$506.69, minus the interest component of \$375.00, or \$131.69. Given this breakdown, the ending balance of the loan at the end of the first month is the starting balance minus the first principal payment, that is, \$100,000 minus \$131.69, or \$99,868.31. For the second payment, interest on the outstanding balance at the end of the first month is $\$99,868.31 \times 4.50\%/12 = \$374.51$; the principal payment is $\$506.69 - \$374.51 = \$132.18$; and the ending balance is $\$99,868.31 - \$132.18 = \$99,736.14$. Subsequent payments are handled analogously.

Appendix A15.1 shows that, using this method of amortizing principal, the ending balance at any time equals the present value of all remaining

TABLE 15.4 Selected Rows of a Mortgage Amortization Table for a 30-Year $100,000 Mortgage at 4.5%. The Total Monthly Payment Is $506.69. Entries Are in Dollars.

Payment Month	Interest Payment	Principal Payment	Ending Balance
			100,000.00
1	375.00	131.69	99,868.31
2	374.51	132.18	99,736.14
3	374.01	132.67	99,603.46
4	373.51	133.17	99,470.29
5	373.01	133.67	99,336.62
60	342.46	164.23	91,157.92
120	301.11	205.58	80,089.43
358	5.66	501.03	1,007.70
359	3.78	502.91	504.79
360	1.89	504.79	0.00

payments discounted at the original interest rate. Applied to the example, the balances at the end of five years, or 60 months (300 months remaining), and at the end of 10 years, or 120 months (240 months remaining), are shown in the table and can be computed, respectively, by the following equations,

$$\$506.6853\frac{12}{4.50\%}\left[1-\frac{1}{\left(1+\frac{4.50\%}{12}\right)^{300}}\right]=\$91,157.92 \tag{15.3}$$

$$\$506.6853\frac{12}{4.50\%}\left[1-\frac{1}{\left(1+\frac{4.50\%}{12}\right)^{240}}\right]=\$80,089.43 \tag{15.4}$$

Figure 15.1 depicts the amortization of this mortgage graphically. For each month, against the left axis, the height of the dark-shaded area represents the principal component of that month's payment; the height of the light-shaded area represents the interest component; and the total height, or sum, equals the fixed total payment of $506.69. The black curve, against the right axis, gives the ending balance or principal outstanding at the end of each month. The figure clearly illustrates the banker's adage that "interest

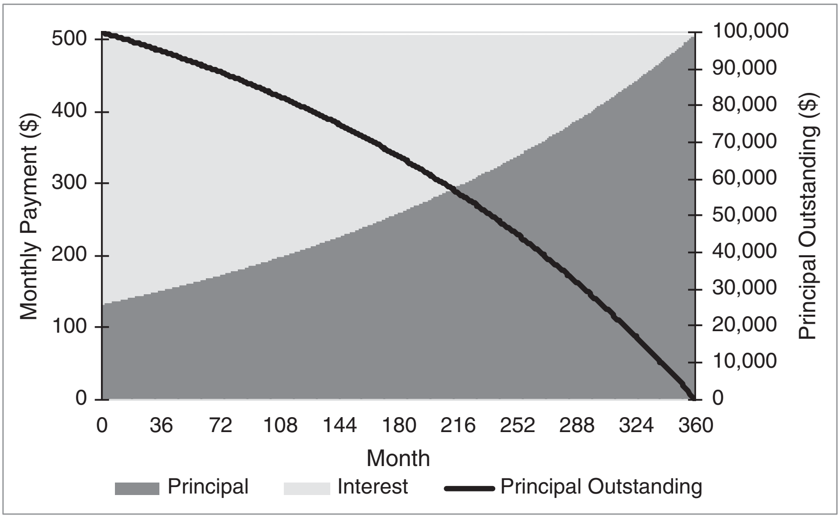

FIGURE 15.1 Amortization of a 30-Year $100,000 Mortgage at 4.5%.

lives off principal," that is, as principal is paid off, interest payments decline. Interest constitutes $375.00/$506.69 or 74.0% of the first monthly payment; $342.46/$506.69 or 67.6% of the 60th payment; $301.11/$506.69 or 59.4% of the 120th payment; etc.; and $1.89/$506.69 or 0.4% of the final payment. While not shown in the figure, the higher the loan rate, the greater the monthly payment, and the greater the proportion of each payment that is dedicated to interest.

15.3 ADJUSTABLE-RATE MORTGAGES

While the overwhelming majority of mortgages are fixed-rate mortgages, *adjustable-rate mortgages* (ARMs) are available as well. The distinguishing feature of an ARM is that the mortgage interest rate can vary over time, but the details of the loan contract can be relatively complex. To illustrate, consider one relatively common variety of ARMs, a 30-year 5/1 *hybrid* ARM with a 2/2/5 *cap structure*.

The phrase "hybrid ARM" means that the mortgage rate is fixed for some number of years before it begins to vary according to some set of rules. In a 5/1 hybrid ARM, for example, i) the mortgage rate is fixed for the first five years; and ii) at the end of that time, and every one year thereafter, the mortgage rate is reset. The new mortgage rate at each reset equals some rate benchmark, like the one-year rate on US Treasury bonds, plus a *gross margin*, which is set and fixed at the initiation of the mortgage.

To put some numbers on this, say that the gross margin is 2.75%, the index is the one-year Treasury rate, and the one-year Treasury rate applied to the reset is 0.25%. In that case, the mortgage rate over the subsequent year is set to 3.00%. The Treasury rate is typically observed with a *lookback*, perhaps 45 days, so that the borrower knows the applicable mortgage rate sometime before it goes into effect.

Once the mortgage rate of an ARM is set, the monthly payments are calculated as in the case of a fixed-rate mortgage. To be more specific, consider the 30-year 5/1 hybrid ARM example at the end of year five. Say that the outstanding balance is $80,000 (computed based on the mortgage rate over the first five years) and that the mortgage rate is to be reset to 3%. The remaining maturity, of course, is 25 years. The mortgage payment is then computed such that the present value of that monthly payment over the next 25 years or 300 months equals $80,000. It is as if a new fixed-rate mortgage for the outstanding balance and the remaining term is being offered at the new mortgage rate.

The 2/2/5 cap structure of the example limits the extent to which the mortgage rate can change over the life of the ARM. The three numbers refer to the *initial adjustment rate cap*, the *periodic adjustment rate cap*, and the *lifetime adjustment rate cap*, respectively. The initial adjustment rate cap of 2% says that the mortgage rate for the sixth year of the mortgage, set at the end of year five, cannot increase or decrease the rate by more than 2% (i.e., a rate of 3% can neither rise above 5% nor fall below 1%). The periodic adjustment rate cap of 2% says that, at the end of any year, the rate cannot increase or decrease by more than 2%. And the lifetime adjustment cap of 5% says that, over the life of the mortgage, the rate cannot increase by more than 5%. Furthermore, the mortgage rate is typically constrained to remain above the gross margin rate.

ARMs typically offer lower initial interest rates than fixed-rate mortgages but expose homeowners to the risk that rates increase. The average loan size of ARMs is significantly higher than that of fixed-rate mortgages, perhaps because wealthier, larger borrowers have more tolerance for interest rate risk. In recent years, the volume of ARMs, as a percentage of all mortgages, has fallen to the single digits. Perhaps, with mortgage rates near historic lows, borrowers perceive the risk of higher rates as exceeding the potential benefit of lower rates. Also, ARMs are still held in some disrepute from the financial crisis of 2007–2009. In the years leading up to the crisis, many people were tempted to buy into the housing boom by the relatively low initial payments of ARMs, accentuated by *teaser rates*, a practice at the time of offering particularly low rates over the first year or two of the ARM. Many homebuyers overextended themselves through these products and faced default as the rates on their ARMs reset higher and the housing market collapsed.

15.4 PREPAYMENTS

The *prepayment option* allows a homeowner to pay off or *prepay* a mortgage, in whole or in part, at any time, by paying the bank the outstanding balance. In the example of Table 15.4, the homeowner may pay the bank \$80,089.43 at the end of 10 years and has no further obligations under the mortgage. But if the mortgage rate has fallen over those 10 years, say to 3.50%, then – using 3.50% instead of 4.50% in Equation (15.4) – the present value of the remaining payments will have risen to \$87,365.60. Hence, in that scenario, the prepayment option is in-the-money: the homeowner may pay the bank \$80,089.43 to extinguish a liability with a present value of \$87,365.60. On the other hand, if the mortgage rate has stayed the same or risen above 4.50% after those 10 years, then the present value of the mortgage liability will be less than \$80,089.43, and the homeowner will have no incentive to prepay.

In practice, of course, most homeowners with mortgages do not have the cash to pay off an existing mortgage. They can take advantage of an in-the-money prepayment option, however, by *refinancing* an existing mortgage. Continuing with the example, the homeowner might borrow \$80,089.43 at the then-prevailing mortgage rate of 3.50%, and then use that cash to prepay the existing 4.50% mortgage. In that case, one way to describe the savings to the homeowner is in terms of the reduction in monthly mortgage payments: using the math given earlier, the monthly payment on a 20-year \$80,089.43 mortgage at 3.50% is \$464.48, which is \$42.21 a month less than the \$506.69 on the existing mortgage. Note that the homeowner is not constrained to refinance with the bank that extended the original mortgage. Another originator might have brought the refinancing opportunity to the attention of the homeowner and might be offering a lower rate or lower transaction costs.

Refinancing activity is usually constrained by the value of the home. Assume, in the example, that the homeowner can economically borrow only 80% of the value of the home; that is, the LTV has to be less than or equal to 80%. In that case, the value of the home when the existing \$100,000 mortgage was extended had to have been greater or equal to \$125,000, and 10 years later, at the time of the \$80,089.43 refinancing, has to be above \$100,112. In fact, if the value of the home after 10 years is greater than that \$100,112, the homeowner might do a *cash-out refinancing*, which means borrowing more money than the outstanding balance of the existing mortgage. In the example, if the home is still worth \$125,000 at the time of the refinancing, the homeowner could borrow \$100,000 in the refinancing mortgage, pay off the existing mortgage with \$80,089.43, and take the residual \$19,910.57 in cash to be used for other purposes. Increasing leverage in this way can, of course, be dangerous. Rising property values in the run up to the financial crisis of 2007–2009 encouraged significant volumes of cash-out refinancings that left many homeowners with negative home equity when home values subsequently plummeted. In fact, in response to that experience, the GSEs currently do not accept cash-out refinancing loans with LTVs greater than 80%.

When interest rates are low, refinancings can generate significant mortgage volumes. Over the most recent period of falling rates, refinancings as a percentage of residential mortgage originations increased from 33% in 2018, to 46% in 2019, and to 65% in the first half of 2020.[7]

Homeowners sometimes exercise prepayment options with *curtailments*, but these are not nearly so significant as refinancings. These prepayments take place mostly when homeowners pay off very old, low-balance mortgages, simply to own a home free and clear of debt. Financial considerations, like the existing mortgage rates relative to market mortgage rates become secondary, lost in the excitement of literal or figurative mortgage-burning parties.

Prepayments occur for other reasons as well, most notably *turnover* and defaults. Turnover refers to the ongoing economic activity of selling and buying homes. Turnover generates prepayments, because homes, which are the collateral for mortgages, cannot be sold until the mortgages are paid off. Disasters that destroy homes also generate turnover prepayments, because mortgages become due in those situations as well.

Defaults refer to cases in which homeowners fail to make contractual mortgage payments. In these cases, for mortgages packaged into MBS, interest and principal are advanced by the servicer and ultimately paid by the insurer. Hence, from the perspective of mortgage investors, defaults are economically equivalent to early payments of principal, rather than as losses to principal invested.

15.5 MORTGAGE POOLS

The simplest form of an MBS is a mortgage *pool*, which is a portfolio of individual mortgages through which payments of interest and principal are passed from borrowers to investors. Borrowers make payments by the first of each month, which servicers pay to investors on the 25th of each month, or on the first business day thereafter.[8]

Table 15.5 presents descriptive statistics as of December 2021 for three 30-year FNMA pools, that is, portfolios of mortgages with approximately

[7]Federal Housing Finance Agency (2021), "What Types of Mortgages Do Fannie Mae and Freddie Mac Acquire?" April 14.
[8]The rest of this chapter ignores this payment delay, but practitioners do account for it, of course, when pricing MBS.

30 years to maturity securitized by FNMA. Consider first the two pools, CA2797 and MA3538, which were both issued in November 2018 and mature 30 years and one month later, in December 2048. Metrics in the table can be shown as "current," that is, as of the report date, in this case December 2021, or as "original," that is, as of the issue date, in this case, November 1, 2018. The table shows the *WAM* or weighted-average maturity of the loans in the pool and the *WALA* or weighted-average loan age, where each loan is weighted by its balance or principal outstanding as a percentage of total pool balance. In the simplest possible case, every loan in a pool would mature in exactly 30 years or 360 months as of the date of issue – giving an original pool WAM of 360 and WALA of zero – and, as of the report date, 37 months later, a current WAM of 360 – 37 = 323 and WALA of 37. Reality is more complex, however, because 30-year pools may be issued containing mortgages ranging in maturity from 181 to 361 months, and many mortgages prepay and drop out of pools between the issue and report dates.[9]

The next row of Table 15.5 gives the *coupons* of the CA2797 and MA3538 pools as 4.5% and 5.0% respectively. The coupon of a pool is the mortgage rate paid to investors. As a pass-through security, pool interest from the borrowers is paid to the investors, but only after subtracting a guarantee fee or *g-fee* for the insurer, in this case FNMA, and a servicing fee for the servicer. The next row in the table, gives the *WAC*, or weighted-average mortgage rate of the loans in the pool. As of the issue date, the WAC of CA2797 was 4.924%, leaving 42.4 basis points worth of fees above the 4.500% coupon paid to investors. The acronym WAC stands, for "weighted-average coupon" and is unfortunate: the word "coupon" most often refers to the rate on a pool paid to investors, while WAC, as just mentioned, is a weighted average of the rates on the underlying loans. In any case, the current WAC of CA2797, at 4.198%, is marginally lower than the original WAC, which means that, on average, loans with marginally higher rates prepaid faster between the issue and report dates. For MA3538, however, the original WAC is slightly less than the current WAC, which means that loans with marginally lower rates prepaid faster.

The next section of Table 15.5 describes the principal amount of the pools, their underlying loans, and experienced prepayments. At issue,

---

[9]By way of illustration, the WAMs and WALA of CA2797 can conceivably be reproduced as follows. Say that, at issue, the pool has 90% of its balance in new loans maturing in 361 months and 10% of its balance in new loans – perhaps used for refinancing – maturing in 351 months. Then assume that, by the report date, all of the 361-month loans prepaid and none of the 351-month loans prepaid. That would leave, as of the report date, a pool with a WAM of 351 – 37 = 314 and a WALA of 37.

TABLE 15.5 Descriptive Statistics for Three 30-Year FNMA Pools, as of December 2021.

Pool	CA2797		MA3538		AI4813	
	Current	Original	Current	Original	Current	Original
Payment Terms						
Issued	11/1/2018		11/1/2018		6/1/2011	
Maturity	12/1/2048		12/1/2048		6/1/2041	
WAM (months)	314	360	319	359	214	359
WALA (months)	37		37		126	
Coupon (%)	4.500		5.000		4.500	
WAC (%)	4.918	4.924	5.772	5.768	4.935	4.942
Loans and Prepayments						
Amount ($mm)	43.654	130.995	105.011	594.633	29.374	266.801
Number	339	943	422	2,050	311	2,037
Avg. Size ($)	128,773	138,913	248,841	290,065	94,451	130,977
1mo CPR (%)	24.6		46.9		21.9	
Credit Quality						
WAOCS		740		678		758
SATO (bps)		−6		78		−6
WAOLTV (%)		79		80		70
Additional Statistics						
Own/2nd/Inv (%)		91/6/3		86/3/11		86/9/5
Buy/Refi (%)		75/25		42/58		53/47
R/C/B (%)		63/25/13		42/46/12		38/58/4
State (%UPB)	TX: 9.0		CA: 20.6		CA 10.6	
	FL: 8.7		TX: 13.5		TX 10.0	
	PA: 6.5		FL: 13.4		FL 6.4	
Pricing						
Price	109.783		108.963		110.309	
OAS	90.1		113.0		86.5	
OAD	3.98		3.18		4.26	

WAM: weighted-average maturity; WALA: weighted-average loan age; WAC: weighted average coupon; CPR: conditional prepayment rate; WAOCS: weighted-average original credit score; SATO: spread at origination; WAOLTV: weighted-average original loan-to-value ratio; Own/2nd/Inv: owner-occupied, second home, investment property; R/C/B: retail, correspondent, broker; UPB: unpaid balance.

CA2797 had 943 loans at an average size or loan principal amount of \$138,913, for a total pool balance of about \$131 million. As of the report date, the number of loans had fallen to 339, and the pool balance to about \$44 million. These numbers reveal a significant rate of prepayment. Using the formulae presented earlier, a \$100,000 30-year mortgage at a rate of 5% amortizes down to an outstanding balance after 37 months of \$95,204. This scheduled amortization can be summarized with a *factor*, or remaining fraction of original principal, of \$95,204/\$100,000 or 0.952. The factor of CA2797 after 37 months, however, is \$43.654/\$130.995 or 0.333. That CA2797 prepaid heavily is not particularly surprising, as mortgage rates fell significantly between November 2018 and December 2021. That the average loan size fell from \$138,913 to \$128,773 indicates that larger loans prepaid more rapidly than smaller loans. This aspect of prepayments is accepted as broadly applicable: homeowners with larger loan amounts are, on average, wealthier and more financially sophisticated and, therefore, more able to take advantage of refinancing opportunities.

The MA3538 pool could reasonably be expected to prepay faster than CA2797. First, the loans underlying MA3538 have higher loans rates, which – subject to a caveat discussed next – increase the incentive to refinance. Second, the average loan size of MA3538 is much higher and, as just mentioned, larger loans tend to prepay faster. And, indeed, MA3538 has prepaid much faster. Its factor of \$105.011/\$594.633 or 0.177 is a lot lower than the 0.333 factor of CA2797.

The next row of the table gives the *CPR*, the *conditional prepayment rate* or the *constant prepayment rate*, over the most recent month. CPR is an annualized measure of the speed of prepayments, which is intricately related to another measure, the SMM or *single monthly mortality* rate. Define SMM_n as the percentage of principal outstanding that is prepaid during month n. Note that prepaid principal does not include scheduled or amortized principal payments, like those described in Table 15.4. CPR then annualizes SMM by assuming that SMM is constant over the year. Under that assumption, the principal not prepaid over one month is $1 - SMM_n$; principal not repaid over 12 months is $(1 - SMM_n)^{12}$, and prepayments over 12 months are given by,

$$CPR_n = 1 - (1 - SMM_n)^{12} \tag{15.5}$$

While CPR is an annualized rate, it is reported monthly based on the prepayments during that month. For example, if a pool prepaid 2% of its outstanding principal above its amortizing principal in a given month, it is prepaying at an annualized speed or CPR given by Equation (15.5), that is, 21.53%. CPR is less than 12 times the monthly rate of 2%, or 24%, because of the compounding in Equation (15.5): the amount remaining each month is not 98% of the original principal balance, but only of the principal balance after some number of months of prepayments. A CPR can be computed over several months as well. Say, for example, that the SMMs for a particular pool over three months were 2%, 2.5%, and 3%. Then the CPR of that pool over that three-month period is,

$$1 - [(1 - 2\%)(1 - 2.5\%)(1 - 3\%)]^4 = 26.21\% \tag{15.6}$$

where raising the product of the three monthly survival rates to the fourth power gives the annual survival rate. Finally, then, returning to Table 15.5, the most recent monthly CPR of MA3538, at 46.9%, is much greater than the 24.6% CPR of CA2797.

The next three rows of the table relate to pool credit quality. The first, *WAOCS* (*weighted-average original credit score*), gives the weighted-average credit or FICO scores of the borrowers with loans in the pool. The second, *SATO* or *spread at origination*, gives the weighted-average spread of the loan rates over a mortgage index rate at the time of origination. By these two metrics, the loans in MA3538 are significantly less creditworthy than those in CA2797. In fact, the 84.4-basis-point difference between the original WACs of the pools is approximately equal to the 84-basis-point difference in their SATOs. This, then, is the caveat in attributing faster prepayments of MA3538 to its higher coupon rate. Because that higher coupon is most likely due to inferior credit, refinancing the loans in MA3538 will likely require a higher rate as well, meaning that the incentive to refinance – which is closely related to the difference between the existing mortgage rate and the available refinancing rate – may not be much greater for MA3538 than for CA2797. Furthermore, lower credit loans are thought to prepay relatively slowly, because homeowners with lower credits might find it more difficult to obtain refinancing loans. In any case, with respect to the other metric of loan quality, *WAOLTV* (*weighted-average original LTV*), the two pools are quite similar, at 79% and 80%.

The next several rows of Table 15.5 report a selection of additional pool statistics. These and other statistics, not shown here, are studied in great detail by market participants in an attempt to determine the speed at which a pool will prepay. The first row of this section of the table gives the percentages of loans for owner-occupied homes, for second homes, and for investment properties. The second gives the percentage of loans used to purchase the home and the percentage used to refinance. The third gives the percentages in each of origination channels described earlier, and the fourth gives the largest percentages of *unpaid balance* by state. Exactly how these statistics predict prepayments is part of the art of prepayment modeling, but some broad outlines are described here. Rates on loans for investment properties are typically higher than for owner-occupied properties, reflecting the greater propensity of investors to walk away from distressed situations. Across mortgages with the same coupon, therefore, investor properties, which have to be refinanced at relatively high rates, face a lower refinancing incentive and, consequently, refinance at lower speeds. With respect to turnover, however, investor properties prepay relatively fast: investors are on the lookout for opportunities to sell properties at a profit, and they do not bear the costs of changing residences to do so. Loans for the purpose of refinancing might prepay faster, because homeowners who have refinanced already might have greater propensity to do so. Brokered originations might refinance faster, because, with no stake in the servicing fees, that is, with no losses from extinguishing existing mortgages, brokers might have the greatest incentive to generate new refinancing originations. This effect is countered, however, by the efforts of originators, mentioned earlier, to recapture refinancings. Finally, loans in some states seem to prepay faster than others, controlling for the other effects discussed so far. These differences across states are usually attributed to differences in the costs of closing mortgages, which are composed of various taxes and fees. For example, average closing costs in New York are among the highest, and prepayments there are particularly slow, while average closing costs in California are relatively low and prepayment speeds there are particularly fast.

15.6 PREPAYMENT MODELING

The complex behavior of prepayments, combined with their importance in mortgage valuation, has given rise to a large industry of building models to predict prepayments as a function of interest rates and other variables. While some highlights are presented in this section, and Appendix A15.2 presents some of the more technical issues encountered, a fuller description is beyond the scope of this chapter.

Prepayment models are often divided into separate modules, one for each driver of prepayments. The most notable drivers, mentioned earlier, are refinancing, turnover, defaults, and curtailments. Traditionally, a module makes a set of assumptions that depend on various parameters. These parameters are then estimated historically, using the vast amount of data available on the prepayment behavior of mortgage pools. Some parameters might also be implied from prevailing market prices of various pools. Most recently, artificial intelligence and machine learning technologies have also been brought to bear on the problem of predicting prepayments.

At the heart of most refinancing modules is an S-curve, which describes refinancings, often in terms of CPR, as a function of some *incentive function*. The incentive function might be the difference between the existing mortgage rate and the current mortgage rate; the ratio of the existing rate to the current rate; the present value of the savings from refinancing; or even the monthly savings from such a refinancing. The S-curve than maps the

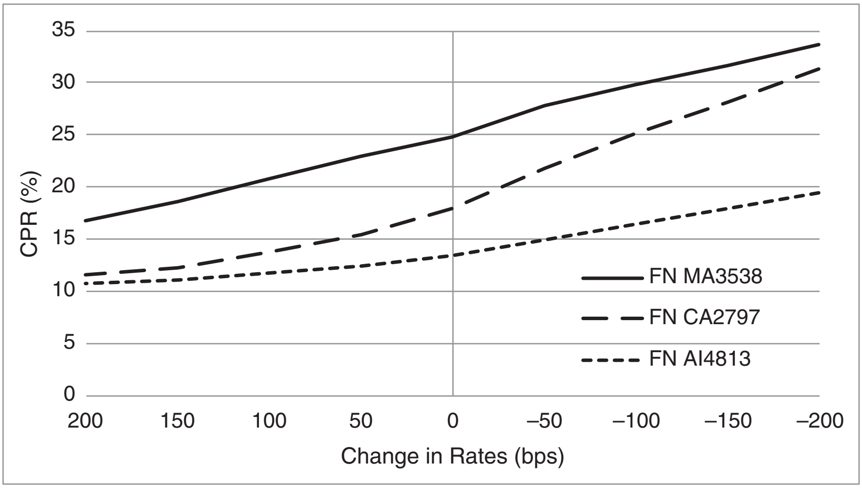

FIGURE 15.2 Refinancing S-Curves for Three Mortgage Pools.

refinancing incentive into a CPR. Figure 15.2 shows the S-curves of a particular model for the three pools summarized in Table 15.5. The horizontal axis here is simply the change in the current mortgage rate. Because existing mortgage rates are fixed, of course, rising current mortgage rates imply less of an incentive to prepay, while falling current mortgage rates imply more of an incentive to prepay.

The name "S-curve" comes from the shape of the functions. For large increases in rates, or, equivalently, very low incentives and high current mortgage rates, CPR is very low, while for large decreases in rates, or very high incentives and low current mortgage rates, CPR is very high. The curves in the figure do not look much like an "S," but the faster a curve moves from low to high CPR, the more the curve looks like something like an "S." In any case, the shape of the S-curve is different for each pool, and can depend on all of the characteristics discussed in the context of Table 15.5 (e.g., average loan size, creditworthiness). According to the models behind Figure 15.2, MA3538 prepays faster than CA2979, but, for a large enough decrease in mortgage rates, the CPR of CA2797 comes close to catching up. Figure 15.3 shows some empirical evidence with which these two S-curves might have been estimated. As the current coupon mortgage rate (which is defined herein) declined dramatically from between 3.50% and 4.00% at the time of origination to less than 1.50% in fall 2020, the CPR of both pools rose dramatically. But MA3538 prepaid faster. As rates stayed low and increased somewhat, prepayments on both remained relatively high, but MA3538, on average, continued to prepay faster. By the way, this figure also shows why refinancing models are not cast as an optimization problem.

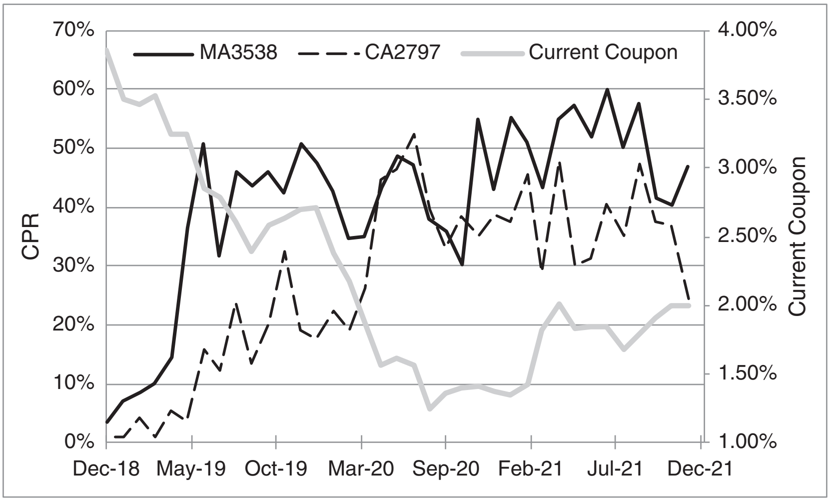

FIGURE 15.3 One-Month CPR for FN MA3538 and FN CA2797 and the 30-Year Current Coupon Rate.

If homeowners exercised their prepayment options with the efficiency of bond issuers exercising their call options, the pools in the figure would have experienced such dramatic prepayments through September 2020 that little to no principal would remain outstanding.

Returning to Figure 15.2, the CPR of AI4813 is much lower than that of the other pools and picks up relatively slowly as rates decline. Contrasting its characteristics with those of CA2797 and MA3538, Table 15.5 reports that AI4813 was issued in June 2011, with a correspondingly shorter current WAM and WALA than the other pools. Its current WAC is essentially the same as that of CA2797. The current average loan size of AI4813 is significantly smaller than that of either of the other pools, while its creditworthiness is marginally higher than that of CA2797. AI4813 is between the other two pools with respect to loan purpose, and the percentage of brokered originations is somewhat lower. The most likely explanation for the lower S-curve of AI4813, therefore, is not these loan characteristics but rather a prepayment phenomenon known as *burnout*.

When a *seasoned* or older pool has been through a period of relatively low rates and experienced significant prepayments, it is most likely that the homeowners with the greatest propensity to prepay have already prepaid and are no longer in the pool. With the remaining loans in the pool less likely to prepay, the pool experiences some degree of burnout. Figure 15.4 illustrates burnout in the case of AI4813. As the mortgage rate fell from about 4.00% at the time of origination to between less than 2.00% and about

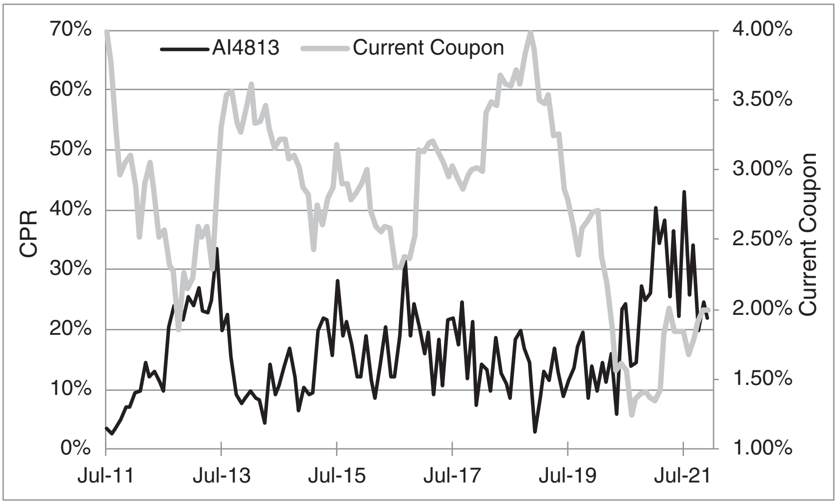

FIGURE 15.4 One-Month CPR for FN AI4813 and the 30-Year Current Coupon Rate.

2.50% from late 2012 through the first half of 2013, CPR increased dramatically, to a peak of about 34%. Rates then rose and fell again to between 2.00% and 2.50%, and CPR reached a similar peak of over 30% in fall 2016. Rates subsequently rose a second time, peaking in November 2018, before starting on a third and more dramatic decline. This time, however, even when rates fell to 2.00%, CPR never rose above 20%. As a burned-out pool, it was not until sometime after mortgage rates has fallen to new lows of below 1.50% that prepayments rose again, this time to a peak above 40%. From a modeling standpoint, burnout introduces the significant complication that the S-curve depends not only on the current characteristics of a pool but also on the history of rates experienced by the pool.

Mortgage professionals might also take the surge of AI4813 prepayments in 2021, as depicted in Figure 15.4, as evidence of a *media effect*. After a precipitous decline in mortgages rates, or after a decline in which mortgage rates reach new lows, media reports and party conversations encourage even those borrowers with low propensities to refinance to do so. Once again, this effect considerably complicates the determination of the S-curve.

Another module of prepayment modeling captures turnover. A common approach to these modules is to start, at issue, from some base rate of turnover, which increases over time according to a *seasoning ramp*. This choice reflects the empirical regularity that homeowners are unlikely to move soon after taking out a mortgage. The exact specification of the seasoning ramp can depend on factors other than mortgage age, most obviously the season, as homeowners are more likely to move at certain times of the year, like soon before schools open rather than soon after. While prepayments due to turnover are for the most part independent of mortgage rates, there is a significant interaction. Borrowers are less likely to move if they have below-market mortgage rates, that is, if they would pay significantly higher mortgage rates after selling their homes and buying new ones. This phenomenon is known as the *lock-in* effect.

As for modules for default, only a few observations are recorded here. Investors in agency MBS experience defaults as prepayments, but defaults are nevertheless tracked and modeled separately, as *involuntary* rather than *voluntary* prepayments. Inputs to these models include credit scores at the time of origination, because credit scores are typically not updated after issuance. While LTVs are not updated either, they can be estimated over time, at the loan level, by using data on home prices in a relevant region. Home prices are important in determining defaults, because a homeowner without cash – but with a home valued in excess of its mortgage balance – can sell the home and pay off the mortgage rather than default.

15.7 MORTGAGE PRICING, SPREADS, AND DURATION

Combining a prepayment model with a risk-neutral model of the term structure, MBS can be computed as the expected discounted value of their cash flows. Because the model price of a given MBS typically does not perfectly match its market price, an *option-adjusted spread* (OAS) is defined as the spread that, when added to the initial, benchmark term structure, results in a model MBS price that equals the market price. OAS is the equivalent of bond spread introduced in Chapter 3 and discussed in Chapter 7. In particular, the OAS can be interpreted as the extra return earned when interest rate risk is correctly hedged by the model with securities that are fairly priced by the model.

A straightforward use of OAS is to buy cheap securities, those with high OAS; to sell rich securities, those with low OAS; and, more aggressively, to engage in relative value trades that simultaneously buy high-OAS and short low-OAS securities. A practical challenge to investing and trading with OAS is model error or misspecification. It is unusual for a model to find that a particular MBS is cheap or rich, while other MBS, with similar characteristics, are fair. More usual is to find that a particular sector, for example, pools with high loan balances, is cheap or rich relative to other MBS. In that case, the success of any trades based on those OAS depends crucially on whether the model is really better than the market at valuing pools with different loan sizes. Another challenge, particularly for relative value trades, is whether OAS are mean reverting. A model may be correctly specified in concluding that a sector is cheap, so that buying and holding MBS in that sector outperforms other sectors over time. But if the sector stays cheap for an extended period of time, a relative value strategy of buying that sector and shorting another might not profit quickly enough to compensate for the use of capital and the costs of financing over time.

Another important use of mortgage pricing models is to estimate the risk sensitivities of MBS. Calculating the duration of the mortgage in Table 15.4 as if its cash flows were fixed, that is, by shifting the discount rate alone, along the lines of Chapter 4, gives a duration of about 11.7. But this fixed cash flow or *static* duration greatly overstates the extent to which the value of a 30-year MBS increases as rates fall. While the value of surviving cash flows does increase as rates fall, all balances prepaid by homeowners are worth only par. To illustrate, Figure 15.5 graphs the price of CA2797 as a function of the five-year Treasury rate from November 30, 2018, to December 21, 2021, excluding the worst of the market turmoil from the pandemic and economic shutdowns, from March 2, 2020, to May 15, 2020.[10] For intermediate and higher rates in the graph, the price of the pool behaves like that of a fixed-rate security: price falls with rates in a close to linear or slightly positively convex manner. But as rates fall, price does not rise monotonically, but levels off at between about 111 and 112 per 100 face amount. Price does not increase further, despite the 4.500%

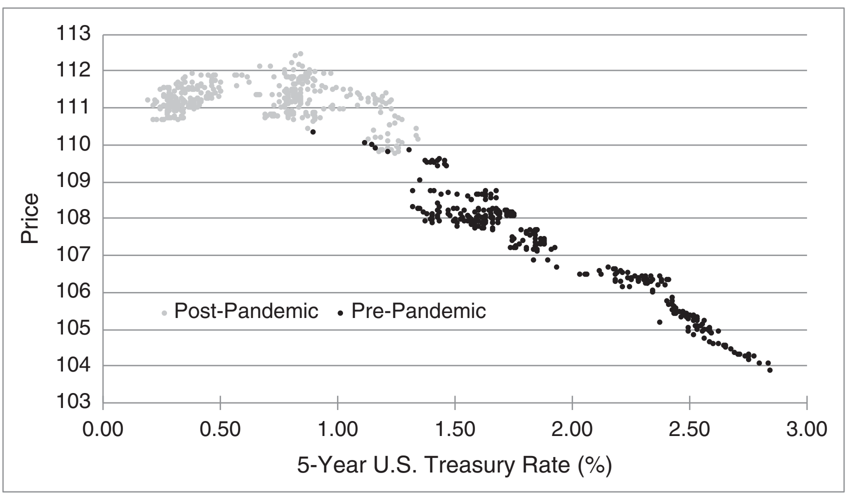

FIGURE 15.5 Prices of FN CA2797 Versus the Five-Year Treasury Rate, from November 30, 2018, to February 28, 2020, and from May 18, 2020, to December 21, 2021.

[10] In the excluded period, mortgage prices fell dramatically – and their OAS relative to Treasuries increased dramatically – because of market stress and increased volatility rather than because of changes in the level of interest rates.

coupon, because of prepayments. But the price does rise significantly about 100, because many homeowners do not prepay, and the present values of their 4.500% cash flows increase significantly in the markedly lower rate environment. Note too that the price–rate curve implicit in Figure 15.5 is negatively convex when rates are below 1.25% or so, that is, as defined in Chapter 4, interest rate sensitivity falls as rates fall.

A mortgage model, with its prepayment modules, can calculate duration meaningfully by shifting both discount factors and cash flows in response to a shift in the benchmark curve. In other words, as benchmark rates increase, discount factors decrease, and prepayments decrease. Along these lines, the percentage change in price scaled for a 100-basis-point decrease in rates gives an *option-adjusted duration* or OAD.[11] Using one particular model, Figure 15.6 graphs the OAD of an index of 30-year FNMA MBS, along with the five-year Treasury rate, over a 10-year period. The OAD in the graph averages 4.3 and ranges from 1.5 to 6.3. As expected, all of these values are significantly less than the static duration of a 30-year 4.5% mortgage, calculated earlier as 11.7. The negative convexity of MBS can be seen clearly from this graph as well. For most time periods, OAD increases as the five-year Treasury rate increases, and OAD decreases as the five-year rate decreases.

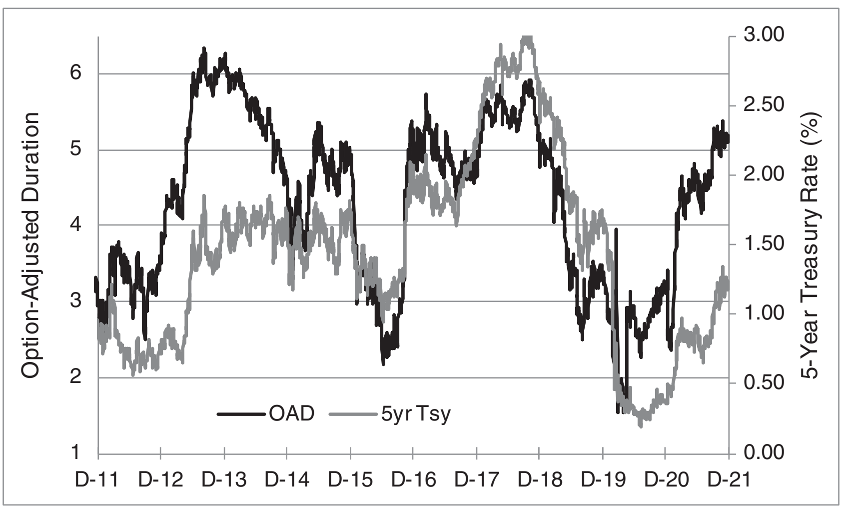

FIGURE 15.6 OAD for a FNMA 30-Year Index and the Five-Year Treasury Rate, December 21, 2011, to December 21, 2021.

[11]As explained in Chapter 4, duration is computed using smaller rate perturbations, but quoted in units that correspond to a percentage change in price for a 100-basis-point decrease in rate.

15.8 TBA AND SPECIFIED POOLS MARKETS

The vast majority of secondary trading of pass-through securities takes place in the very liquid *TBA* market, while the rest takes place in the less liquid *specified pools market*. The TBA market is a forward market in securitized mortgage pools, in which the seller has a delivery option. Table 15.6 shows a snapshot of TBA prices for UMBS 30-year pools from late December 2021. The rows correspond to different delivery months, the columns correspond to different mortgage pool coupons, and the entries give the corresponding forward prices. Most of the liquidity in the TBA market is concentrated in a few coupons, although a fuller *stack* of coupons do trade. Most of the liquidity is also in the *front* or first contract and the second contract, but later delivery months trade as well. The front contract with the greatest price less than par is called the *current contract*. The *current coupon*, or *current coupon rate*, which is often quoted as a description of the prevailing mortgage rate, is defined as the coupon that would give a TBA contract priced at par, calculated by interpolating the coupon of the current contract and the coupon of the front-month contract with a price just above par. From Table 15.6, the current contract is the January 2.0% TBA, and the current coupon, calculated from the coupons and prices of the January 2.0% and January 2.5% TBAs, is about 2.11%.

A seller of $1 million face amount of the January UMBS 30-year 2.5% TBA commits to sell $1 million principal amount of a 30-year 2.5% UMBS pool (i.e., one issued by FNMA or FHLMC) at 101–28 for settlement or delivery in January 2022. The TBA buyer commits to buy the pool delivered by the seller. The exact settlement date is published well in advance, and is always in the stated delivery month, but varies over time and by product.[12] With respect to Table 15.6, the settlement dates for the three months are January 13, 2022, February 14, 2022, and March 14, 2022.

TABLE 15.6 Bid Prices for UMBS 30-Year TBAs as of December 30, 2021.

Coupon	1.5%	2.0%	2.5%	3.0%
Jan	96–13	99–15+	101–28	103–13
Feb	96–08	99–08+	101–20	103–08+
Mar	96–03	99–02	101–12+	103–03+

Prices are given as a handle and a number of ticks or 32nds; for example, 101–12+ is a price of 101 + 12.5/32, or 101.390625.

[12] "Class A" settlement dates apply to 30-year UMBS TBAs; "Class B" to 15-year UMBS and GNMA; "Class C" to 30-year GNMA; and "Class D" to an assortment of other products.

The seller of a TBA has a delivery option in the sense of being able to deliver any pool that falls within the general parameters just described.[13] For this purpose, a "30-year" pool is defined as one with a remaining maturity of between 15 years and one month to 30 years and one month. The name TBA is an acronym for to be announced, and comes from the seller's option to choose the pool to be delivered after the trade date. More specifically, the seller must notify the buyer of the exact pool to be delivered on the *notification date* or *48-hour day*, which is two business days before the settlement date. For the contracts in the table, the notification dates are January 11, 2022, February 10, 2022, and March 10, 2022. Note, by the way, that market practitioners always say TBA, never *to be announced*.

Chapters 11 and 14 describe how the pricing of Treasury note and bond futures and of credit default swaps (CDS) reflect seller choices to deliver the least valuable or "cheapest-to-deliver" (CTD) securities. Similarly, TBA prices reflect seller choices to deliver the least valuable eligible mortgage pools, which are typically those that refinance with the greatest speed and least predictability. The analogy to Treasury futures and CDS is not perfect, however. There is typically one security that is clearly CTD into a Treasury futures or CDS contract, while many mortgage pools might be delivered into a TBA. In fact, an important part of trading in this market is forecasting the pools likely to be delivered into TBAs given the supply and characteristics of deliverable pools outstanding. In the case of the January 2.5% TBA, for example, one dealer forecasts the delivery of pools with a WAC of 3.35%, a WALA of three months, and an average loan size of $340,000. While much analysis is needed to explain the precision of this prediction, it makes sense that pools with these characteristics have less value than other 30-year 2.5% pools: higher WACs typically prepay faster (with the credit score caveat mentioned previously); relatively new pools typically prepay faster and have more prepayment uncertainty than seasoned, somewhat burned-out pools; and pools with high average loan balances prepay faster as well. (Note that an average loan balance of $340,000 is large relative to the averages presented in Table 15.3.)

The other secondary market for pass-through trading is the specified pools market. Because the least valuable pools are delivered into TBAs, which is reflected in TBA prices, sellers with more valuable collateral, along with buyers wanting more valuable collateral, prefer to trade particular or specified pools. Prices of pools traded in this market are quoted as a *pay-up* relative to the comparable TBA. Table 15.7 shows a sample of pay-ups for UMBS 30-year specified pools as of December 2021. The first four rows are pools with loan balances in particular ranges. Pools with a coupon

[13] Some TBA trades include stipulations or *stips* that add restrictions on deliverable pools.

TABLE 15.7 Representative Pay-ups for Selected Specific UMBS 30-Year Pools, in Ticks (32nds), as of December 2021.

Specified Pool Description	Coupon 2.0%	2.5%	3.0%
Average Loan Balance			
Less than $85,000	36	58	109
Between $85,000 and $110,000	31	49	88
Between $125,000 and $150,000	23	37	71
Between $200,000 and $225,000	10	16	44
100% NY Loans	9	28	75
100% Investor Property	3	14	34
Jumbo	−14	−18	−38

of 2.5% and average loan balances between $125,000 and $150,000, for example, sell at a pay-up of 37 ticks, that is, for 37/32 or 1.15625 over the price of the UMBS 30-year 2.5% front contract TBA. The pay-ups in each of the three coupon columns decrease with average loan balance, consistent with the empirical regularity that pools with higher loan balances refinance faster. That pools with average loan balances between $200,000 and $250,000 command a pay-up reflects the expectation, mentioned in the previous paragraph, that CTD pools have an even larger average loan balance of $340,000. For any given balance range, the pay-up increases with coupon: relatively low loan balances and slower refinancing speeds are more valuable the higher the coupon, that is, the greater the incentive to prepay and the greater the loss in value to investors as a result of refinancings.

The next row of Table 15.7 gives the pay-ups for pools made up of loans in New York State, which, as mentioned earlier, refinance slowly relative to loans in other states. And, once again, those lower refinancing speeds are worth more the greater the coupon. The next row is for pools composed of loans for investor property. As discussed already, investor properties face higher refinancing rates, and, therefore, refinance more slowly, resulting in positive pay-ups that increase with coupon.

The last row of the table gives pay-ups for pools containing jumbo loans, that is, loans of size greater than agency conforming limits. This row differs from others in that jumbo pools are not deliverable into TBAs. In any case, jumbo pools trade at a negative pay-up, that is, at lower prices than TBAs, because these extremely large loans are expected to prepay faster than any conforming pool delivered into TBAs.

Table 15.7 gives only a sample of specified pools. Other specified pools, with other defining characteristics, trade as well, like low FICO loans, high

SATO loans, and loans serviced by certain banks. As for this last category, certain origination channels and servicing arrangements are more aggressive than others in encouraging homeowners to refinance. As a result, a pool of loans in less aggressive settings, like those serviced by certain banks, are expected to prepay more slowly and, therefore, command a pay-up in the specified pools market.

15.9 RISK FACTORS AND HEDGING AGENCY MBS

Participants in agency MBS face and have to manage a number of risks: interest rate risk, which includes convexity and volatility risks; mortgage spread risk; prepayment risk; and credit risk. Interest rate risk is discussed extensively in earlier chapters of the book. Mortgages are clearly exposed to the risk that rates rise and fall, and, as described before, exhibit negative convexity. Furthermore, negatively convex positions are also short interest rate volatility.[14]

Mortgage spread risk refers to the risk of changes in the spread between mortgage rates and benchmark rates, that is, Treasury or swap rates. If an MBS is hedged by selling Treasury futures and then mortgage spreads increase, the MBS falls in value relative to the Treasury hedge and the overall position loses money. If, on the other hand, one MBS is hedged by another, perhaps a TBA, then changes in value due to mortgage spreads can offset. Prepayment risk, following earlier discussions, is the risk that prepayments are higher than anticipated for premium mortgages or lower than anticipated for discount mortgages. Finally, the credit risk from agency MBS, the risk of interest and principal loss due to homeowner default, is borne by the agencies that guarantee the underlying mortgages or the MBS.

Various groups of market participants face these risks in different permutations. Originators, who make mortgage loans in order to sell them into MBS, are exposed to *pipeline* risk. Originators offer firm rates to borrowers, called *rate locks*, but it then takes time to close or complete the mortgage deal. A possible hedge against the risk that rates rise before the mortgage can be sold is to short Treasuries or pay fixed in swaps, with hedge ratios calculated as OADs to account for prepayment risk. But this strategy has three problems. First, borrowers can walk away from rate locks at their discretion. While a deal might not close for many reasons, borrowers do tend to abandon offers when rates subsequently decline, while they tend to close on mortgages when rates subsequently increase. Accurate hedging of pipeline risk, therefore, has to account for the exercise of this borrower option. Second, because mortgages are negatively convex, hedging them with

[14]The discussions about interest rate risk in this section rely heavily on the material about convexity from Chapter 4 and material about swaptions in Chapter 16.

positively convex Treasuries or swaps leaves an overall negatively convex and short volatility position. One solution is to buy short-term payer swaptions to manage convexity and to buy longer-term swaption straddles to hedge the volatility.[15] Third, because mortgage values and this borrower option depend on prevailing mortgage rates, rather than on Treasury or swap rates, hedges with Treasuries or swaps are subject to mortgage spread risk.

An alternative hedge of originator pipeline risk is to sell TBAs rather than Treasuries or swaps. Prepayment models are still needed to calculate hedge ratios, and the problem of modeling rate-lock optionality remain. But because TBAs are also negatively convex and short volatility, selling them can offset the negative convexity and short volatility properties of the mortgages being originated. In that case, swaptions are necessary only to address any residual exposures. Furthermore, because both TBAs and the mortgages being originated clearly depend on mortgage spreads, that risk can be mitigated by the TBA hedge as well.

Originators that hedge pipeline risk by selling TBAs have another decision to make when the originated pools are ready for sale. If the pools are significantly more valuable than TBA prices, the originator can buy back the TBA hedge and sell the pools in the specified pool market. If the pools are about as valuable as TBA prices, the originator can deliver the pools into the short TBA hedges. And if the pools are worth somewhat more than TBA prices, there is a trade-off of value and liquidity: the pools may be worth somewhat more if sold as specified pools, but they can be sold with more liquidity in the TBA market.[16]

Mortgage servicers lose revenue when mortgage rates decline and homeowners refinance. This effect tends to dominate the increase in value of the remaining servicing flows due to discounting at lower rates. Mortgage servicing rights, therefore, usually have significant negative duration. They are also negatively convex. Their interest rate sensitivity increases as rates increase: from very negative at low rates, when homeowners are prepaying fast, to less negative at high rates, when homeowners are prepaying slowly. One possible hedge of servicing revenue is to buy Treasuries or receive fixed in swaps. The determination of hedge ratios requires a prepayment model, of course, but the resulting, overall position is negatively convex, short volatility, and subject to mortgage spread risk. The alternative of hedging by buying TBAs is not as appealing a solution as it is for originators. True, buying TBAs across

[15]Swaptions are discussed in Chapter 16. Like any option, a swaption has a positively convex price–rate profile, which can be used to manage the convexity of other positions. Also, like any options, swaptions increase in value with volatility. Shorter-term swaptions, however, are less exposed to volatility than long-term options.

[16]An originator can also sell pools into other securitizations, like CMOs, discussed in a later section.

the stack of coupons can increase the effectiveness of prepayment hedging. And true, buying TBAs against servicing rights hedges mortgage spread risk. But with both mortgage servicing rights and long TBA positions negatively convex and short volatility, the overall "hedged" position is even more negatively convex and short volatility. If that result is unacceptable from a risk management perspective, swaptions can be bought – at a cost – to flatten out the exposure.

As mentioned earlier in the chapter, some servicers are also originators, and servicing and origination businesses, at least to some extent, hedge each other. It might be cheapest, at least in theory, for entities with both businesses to rely on this *operational hedge* and hedge only residual risk with Treasuries or swaps, TBAs, and swaptions. In practice, however, organizational efficiencies might lead each business to hedge its risk independently. In that case, a middle ground might be to save on transaction costs by crossing hedging trades internally.

The various market participants that invest in mortgages can decide which risks to bear and which risks to hedge. A mortgage exchange-traded fund is likely to buy MBS and take all of the associated risks, in the expectation of earning commensurate return. A more actively managed mortgage portfolio might hedge out interest rate risk with Treasuries or swaps, and even negative convexity and volatility with swaptions, but keep mortgage spread risk and prepayment risk, possibly to earn appropriate compensation for bearing those risks over time, or possibly to take advantage of perceived skill in timing exposures to those risks. Finally, a relative value mortgage desk might bet on prepayment speeds alone, essentially depending on its prepayment model outperforming the market consensus and implied pricing. This strategy is most likely implemented by buying relatively cheap specified pools and selling relatively rich TBAs, or *vice versa*, or by trading one TBA coupon against another in the stack.

Last but not least are the agencies. They manage the default risk arising from their guarantees in two ways. First, they collect g-fees that are designed to be sufficient, in aggregate, to make up for experienced losses. Second, agencies sell off some of their default risk to the private sector through credit risk transfer securities, which are discussed presently. The extent to which the agencies bear other mortgage risks depends on the method by which the MBS are issued. In *lender swap transactions*, which are the most common, an agency exchanges mortgage loans from originators for MBS made up of those loans. The loans move directly into a trust backing the MBS without ever residing on the agency's balance sheet. Therefore, in this kind of issuance, all risks (other than default) remain with the originator. In *portfolio securitization transactions*, by contrast, an agency buys mortgages for its own portfolio, packages them into MBS, and sells those MBS in the secondary market. In this method of issuance, the agency bears interest rate risk, prepayment risk, and mortgage spread risk from the time the mortgages are purchased until the time they are sold through MBS. Along the lines of the earlier discussion about private originators, an agency can hedge these risks by selling Treasuries or swaps, or by selling TBAs, and by supplementing these hedges with the purchases of swaptions. Also, like other originators, agencies can decide whether to buy back TBA hedges and sell originated pools in the specified pools market or to deliver originated pools into their short TBA hedges.

This section concludes with a further observation about the TBA and specified pools markets. On the one hand, the TBA delivery option means that TBA prices track those of the least valuable pools. But on the other hand, from a primary market that produces a huge number of distinct pools, the TBA delivery option creates extremely liquid secondary market trading in a handful of TBA contracts. This liquidity allows market participants to hedge their specified pool exposures with TBAs; that is, the liquidity of the TBA market easily compensates for any basis risk arising from differences between changes in specified pool prices and changes in TBA prices. Furthermore, the liquidity of the TBA market is so attractive that many pools are, in fact, physically delivered – and thus sold – through TBA contracts, despite relatively low TBA prices due to the delivery option, and despite the general preference of derivatives users to unwind derivatives hedges and to trade physical or cash product with their regular business counterparties.

15.10 DOLLAR ROLLS

Because TBA liquidity is concentrated in contracts that mature in one of the next several months, market participants often need to roll their positions in maturing contracts to contracts that mature at later dates. For example, originators and investors that are short January 2022 TBAs to hedge exposures but wish to keep those short positions past the January settlement date, can *buy the roll*, that is, buy (back) the front month contract (January) and sell contracts expiring in later months, perhaps in February. Similarly, servicers and investors who are long January 2022 TBAs, but wish to maintain those longs past January, can *sell the roll*, that is, sell (out of) the front month contract (January) and buy contracts expiring in later months. These TBA rolls are also called *dollar rolls*.

A natural question that arises in the context of dollar rolls is whether TBA prices are fair relative to one another. Because TBAs are forward contracts, several of the insights of Chapters 10 and 11 apply here. First, a long TBA position is similar to a long position in the underlying pool and a short repo position (i.e., borrowing money on the collateral of the pool). Second, TBA prices will exhibit forward drops so long as the carry on the pool exceeds the repo rate. This is illustrated, in fact, in Table 15.6. The repo rate at the time is 10 basis points; all the TBA coupons shown significantly exceed that repo rate; and TBA prices are lower for later expiration months. Third, just as each TBA has implicit financing from the trade date to settlement, a TBA roll embeds implicit financing from the expiration of the front contract to the expiration of the deferred contract.

Market practitioners calculate the *value of a roll* and the *implied financing rate* or *breakeven financing rate* of a roll along the lines of the cash-and-carry arbitrage for forward contracts on coupon bonds discussed in Chapter 11. In that context, however, the forward contract is on a particular bond and that bond's coupon is known with certainty. In the TBA context, because of the delivery option, the forward contract is on an unknown pool. Furthermore, because of prepayments, the cash flows of pools are not known with certainty. An industry standard *roll analysis* puts these issues aside, essentially assuming that rolls are perfectly equivalent to mortgage repo. After presenting this analysis, the text returns to the differences between TBA and repo financing.

Consider an investor who holds \$1,000,000 of a 30-year 2.5% pool in early January 2022. Assume that TBA prices are as in Table 15.6 and that the mortgage repo rate of 10 basis points is the relevant risk-free rate. Also, consistent with the discussion in the previous paragraph, assume for the purposes of this analysis that the February cash flow from the pool (payable to investors on February 25) is known and equal to \$10,850.00, which includes \$8,800.00 of principal (both scheduled payments and prepayments).

The investor holding the pool is considering two strategies, both of which leave the investor with the same outstanding balance of the pool on 2/14: i) hold the pool to February 14, the settlement date of the February TBA, at which point the principal outstanding balance is \$1,000,000 minus \$8,800.00, or \$991,200; ii) sell the pool through the January TBA for 101–28 (101.875) and buy back \$991,200 of the pool through the February TBA at 101-20 (101.625). The assumption that the investor will get back the same pool that had been sold in January is again consistent with the simplifications of this analysis. Figure 15.7 summarizes these two strategies: holding the pool, below the line, and ii) rolling, above the line.

The cash flows from holding the pool are very simple. The February payment of \$10,850.00 is received on 2/25, but, for comparison purposes, is discounted 11 days, back to 2/14, for a present value of $\$10,850.00/(1 + 11 \times 0.10\%/360) = \$10,849.67$.

The cash flows from rolling are as follows. The pool is sold through the January TBA for flat proceeds of $\$1{,}000{,}000 \times 101.875 = \$1{,}018{,}750.00$, plus 12 days of accrued interest, that is, $\$1{,}000{,}000 \times 2.50\% \times 12/360 = 833.33$, for total proceeds of \$1,019,583.33. These proceeds are invested at the repo rate from 1/13 to 2/14, that is, for 32 days, to earn interest of $\$1{,}019{,}583.33 \times 0.10\% \times 32/360 = \$90.63$. The next part of this strategy is to buy \$991,200.00 face amount of the pool through the February TBA at

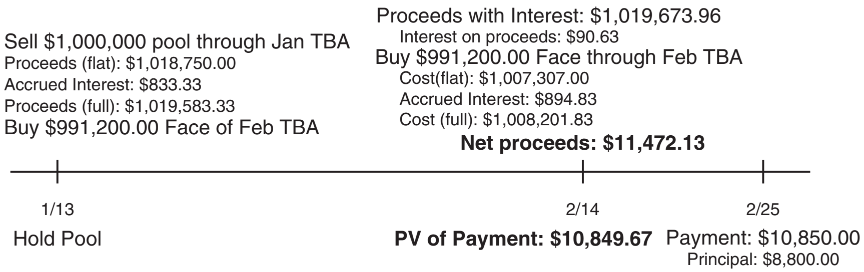

FIGURE 15.7 Dollar Roll Example, UMBS 30-Year 2.5% Jan–Feb TBAs, as of January 2022.

a flat cost of $\$991,200.00 \times 101.625\% = \$1,007,307.00$, plus accrued interest of 13 days of $\$991,200.00 \times 2.50\% \times 13/360 = \$894.83$, for a total cost of $1,008,201.83. Finally, subtracting this cost from the invested proceeds of the January sale leaves a net of $11,472.13.

Putting the pieces together, both the hold and roll strategies leave the investor with $991,200 of the pools as of 2/14. But the roll strategy also leaves the investor with net proceeds of $11,472.13 on 2/14, while the hold strategy leaves the investor with $10,849.67. The value of the roll, defined as the difference between these two amounts, is $622.46. Hence, according to this analysis, the January TBA is rich relative to the February TBA, or, in other words, investors are willing to pay up to have the 2.5% pools now.

The advantage of the roll is often expressed as an implied or breakeven financing rate, which, replacing the repo rate, sets the value of the roll to zero. In this example, the implied financing rate is −58.4 basis points. The difference between the actual repo rate of 10 basis points and this implied rate, or about 68 basis points, is known as the *specialness* of the pools. From a rates perspective, owners of the 2.5% pools can earn 68 basis points over repo by giving up their pools for a month.

While the numbers in this example are rounded for easy reading, the implied financing rate of −58 basis points is consistent with the analysis circulated by market professionals at the time. The discussion circles back then, to the simplifying assumptions behind the standard dollar roll analysis. First, the assumption that the same pool will be returned to the owner in February overstates the desirability of the roll to the owner of the pool: the TBA delivery option opens the roll seller to the risk of receiving a pool of lower value. Second, the prepayment assumption is critical in the calculation of the breakeven rate. Slower prepayments make holding the pool more valuable, or equivalently, reduce the value of the roll and increase the implied financing rate. Third, the assumption of known prepayments overstates the desirability of holding the pool, or equivalently, understates the value of the roll, because prepayments are, in fact, uncertain, and prepayment risk is borne by the holder of the pool.

In summary, financing in the mortgage market is available directly, through repo, and indirectly, through TBAs. The liquidity of the TBA market is a factor in its favor. Repo financing keeps the prepayment risk of intermediate payments with the owner of the collateral, while TBA positions avoid that risk. Repo financing keeps control of the collateral; long TBA positions give up control of collateral; and short TBA positions are long a delivery option.

15.11 OTHER MBS

This section briefly describes MBS other than individual pools. As mentioned earlier, because pass-through securities are relatively large, diversified, and insured portfolios of mortgages, they are owned by a broader investor base and trade with more liquidity than individual mortgages. Taking this another step, the agencies combine pools into larger MBS, to create even more diversified and, if successful, more liquid securities. These conglomerated MBS with FNMA collateral are called *Megas*; with FHLMC collateral, *Giants*; and with collateral from both agencies, *Supers*.

Another type of MBS of significant size are *REMICs* (*Real Estate Mortgage Investment Conduits*) or *CMOs* (*Collateralized Mortgage Obligations*). While pass-through MBS have proved popular with many investors, others are less comfortable with prepayment risk, that is, with securities whose duration can fluctuate widely, as illustrated in Figure 15.6. Some CMO and REMIC classes or tranches address this problem. *SEQs* (*sequential pay classes*), for example, are structured as follows. The cash flows from an underlying portfolio of pools are passed through to a number of distinct SEQ classes. All classes receive monthly interest payments, but principal payments from the underlying pools are directed first to pay off the A class. Once the principal amount of the A class is fully paid off, principal payments from the pools are directed to the B class, and so forth. While SEQ classes still bear significant prepayment risk, their effective maturities can be bucketed with somewhat more granularity than standard pass-through MBS.

PACs, or *planned amortization classes*, are further along in the spectrum of creating more complicated structures so as to create some securities with less prepayment risk. Once again, the cash flows of an underlying portfolio of pools are divided across a number of classes. PAC classes, however, promise fixed schedules of principal payments so long as prepayment rates stay within set ranges. Making this possible are *support* or *companion* classes that absorb prepayment fluctuations within these ranges. If prepayments are relatively fast, then the support classes get paid principal quickly, enabling the PAC classes to maintain their fixed payment schedule. Conversely, if prepayments are relatively slow, the support classes absorb the shock by being paid principal particularly slowly. If prepayment rates move outside the set ranges, however, the PAC bonds will experience faster or slower principal payments as well. Similar to PACs are *TACs* (*targeted amortization classes*), which protect only against faster prepayments, and which are usually included in structures that include PACs as well. The issuance of CMOs offering PACs and TACs is viable, because their classes can be sold for more than the underlying pools. In other words, some investors are willing to pay up enough for MBS with prepayment protection so as to more than compensate the investors who buy and take on the amplified risks of the support classes.

Interest-only (IO) and *principal*-only (PO) strips are another well-known set of MBS products. The cash flows of an underlying portfolio of pools are directed into these two classes by a simple rule: all interest payments are paid to the IOs and all principal payments to the POs. The price sensitivities of these strips are larger than many other fixed income products. As rates fall and prepayments increase, the cash flows of the IOs begin to vanish. This loss of value more than offsets any gains from higher discount factors on remaining cash flows, and IO prices decline. Hence IOs, like mortgage servicing rights, have negative durations at low levels of rates. Conversely, POs gain dramatically in value as rates decline. They have the relatively high interest rate sensitivity of zero coupon bonds, with the added effect of early repayment of principal when rates are low.

15.12 CREDIT RISK TRANSFER SECURITIES

One reform following the 2008 government rescue of the GSEs was to have the GSEs sell some of the mortgage default risk from their MBS guarantees to the private sector. To this end, GSEs started to offer *credit risk transfer* (CRT) securities to the public in 2013, called *Connecticut Avenue Securities* (CAS) at FNMA and *Structural Agency Credit Risk* (STACR) securities at FHLMC.[17] The basic idea is that investors buy bonds at par; earn a spread over a short-term rate; and, in exchange for that spread, bear some of the default risk on a *reference pool* of mortgages. More specifically, as defaults on that pool reach certain thresholds, investors lose some or even all of the principal on their bonds. A somewhat simplified explanation of how CRTs work is presented in this section, with reference to a recent offering, CAS 2020-R01, summarized in Table 15.8. The notes actually offered to the public through CAS 2020-R01 are highlighted in bold, namely, the M1, M2, and

[17]The GSEs have additional programs, not discussed in the text, through which they sell off default risk to insurance companies, called *credit insurance risk transfer (CIRT)* at FNMA and *agency credit insurance structure (ACIS)* at FHLMC.

TABLE 15.8 Connecticut Avenue Securities (CAS) 2020-R01, Issued in January 2020 and Maturing in January 2040. Spread Is Over One-month LIBOR.

Reference-Pool Balance ($mm):				**28,996**		
Class	**Notes ($mm)**	**FNMA ($mm)**	**Credit Support (%)**	**Thickness (%)**	**Rating (Fitch)**	**Spread (bps)**
A		27,851	3.95			
M1	**303**	16	**2.85**	**1.10**	**BBB-**	**80**
M2	**523**	28	**0.95**	**1.90**	**B**	**205**
B1	**207**	11	**0.20**	**0.75**	**NR**	**325**
B2		58	0.00	0.20		
Total	1,033					

B1 classes. The A and B2 classes, along with the M1, M2, and B1 amounts in the FNMA column, are fictional: they are included in the description of CRTs only to explain how the cash flows of the actual notes are determined.

The first step in creating a CRT is to select a group of mortgage loans for inclusion in the reference pool. All candidate loans are conforming but are further scrutinized to exclude those with "underwriting defects" and "performance deficiencies." This screening is one of several ways the GSEs seek to assure CRT investors that they are taking representative mortgage default risk, rather than default risk of a particularly weak group of loans. In any case, the reference pool chosen for CAS 2020-R01, at the time of issue, had over 100,000 loans and a total balance of about $29 billion.

At issuance, investors buy a total of $1.033 billion principal amount of notes at par, $303 million in the M1 class, $523 million in the M2 class, and $207 million in the B1 class. The proceeds from this sale are deposited into a trust that earns enough to pay investors a short-term rate on their principal, here one-month LIBOR. In addition, the GSE commits to contribute some of the g-fees it earns to pay the CRT investors a spread over that short-term rate, in particular, 80, 205, and 325 basis points on the M1, M2, and B1 classes, respectively. In exchange for that spread, the investors bear some of the default risk of the reference pool of mortgages, in a manner to be described presently.

The return of principal to CRT noteholders depends on the principal cash flows of the reference pool. It is important to emphasize, however, that cash flows from the reference pool do not, in any sense, flow to CRT investors. The owner of each loan in the reference pool, whether a GSE or some MBS investor, receives all of the interest and principal cash flows from that loan. The interest and principal paid to CRT investors comes from the CRT trust, with the interest supplemented by the GSE contributions mentioned in the previous paragraph.

As the reference pool makes voluntary principal payments (i.e., scheduled principal payments and prepayments that are not due to default), principal is returned to the CRT notes. The A or senior class receives principal in proportion to its outstanding balance relative to the remaining balance of the reference pool. At issuance, for example, the A class is $27,851/$28,996, or about 96% of the outstanding reference-pool balance. Therefore, the A class is allocated 96% of the voluntary principal payments from the reference pool. Of course, as mentioned earlier, the A class does not really exist, and is just used for calculating cash flows due to the actual notes. But, along those lines, the principal balance of the A class is reduced by these fictional returns of principal. The remainder of principal payments from the reference pool, or about 4% at the time of issuance, is allocated to the M1, M2, B1, and B2 or subordinate classes, in order of seniority. This means that, so long as the M1 class has an outstanding principal balance, all of the principal payments to the subordinate classes are allocated to M1. When the M1 notes are completely paid off, these principal payments are allocated to M2 and so on. Note that the principal payments to the M1, M2, and B1 classes are real. The calculation of the payments are as described, based on the principal payments of the reference pool, but the payments themselves come from the CRT trust into which the original sale proceeds were deposited.

The text now turns to the crux of the structure, namely, the allocation of default losses in the reference pool to CRT note investors. Once again, to assure CRT investors, the issuing agency holds the first-loss tranche, which, for CAS 2020-R01, is the fictional B2 class. According to Table 15.8, FNMA held $58 million of this class, which is assigned losses starting from a *credit support* of 0% to a total *thickness* of 0.20% of the original $28.996 billion reference-pool balance, which is also $58 million.[18] If the reference pool suffers $1 million of losses, the principal of the B2 class falls from its original $58 million to $57 million. If the pool experiences another $5 million of losses, another $5 million of B2 principal is written down, to $52 million, and so forth. If cumulative losses ever reach $58 million, the B2 class is written down to zero.[19]

If losses are large enough to burn through the first-loss tranche, the real CRT investors start to lose principal. According to Table 15.8, the original principal of the B1 class was $218 million, of which 95% or $207 million was held by private investors, and 5% or $11 million was assigned to

[18]The assignment of losses in CRTs is similar to that of collateralized loan obligations, described in Chapter 14, but the nomenclature is different. The "attachment point" of CLOs corresponds to the credit support percentage here, and the "detachment point" of CLOs equals the sum of the credit support and thickness percentages here.

[19]The text abstracts from the fact that recoveries in the event of default are uncertain. Therefore, losses assigned to tranches can be subsequently reversed if recoveries are greater than anticipated.

FNMA. FNMA holds 5% of the B1 class, as well as 5% of the M1 and M2 classes, as "skin-in-the-game," to assure investors of each class that their interests and FNMAs are aligned. These 5% holdings of each class are examples of *vertical slices*, because they are treated just as investor holdings in the same class. *Horizontal slices*, by contrast, are treated better or worse than other classes, like the B2 and A classes.

With credit support of 0.2%, B1 does not suffer any principal write-downs until reference-pool losses exceed 0.2% or $58 million, or, equivalently, until the B2 class has been wiped out. Any further losses, however, are for the most part borne by B1 noteholders. If the reference pool experiences $1 million of losses beyond $58 million, 95% of those losses or $950,000 result in a write-down of B1 principal from the original $207 million to $206,050,000, and 5% or $50,000 of those losses result in a write-down of FNMA's fictional B1 principal from $11 million to $10,950,000. Operationally, the CRT trust sends $950,000 to FNMA, which uses those funds either to compensate MBS investors, or, if the loans are held in its own portfolio, to avoid suffering losses on its own account.

The B1 class is wiped out itself if losses beyond $58 million equal the class thickness of 0.75% of the original reference pool, that is, 0.75% times $28.996 billion, or $218 million. In that case, further losses hit first the M2 and then the M1 classes, in the same way as just described for the more junior classes. All in all, FNMA's B2 class bears the first 0.2% of losses; the B1, M2, and M1 classes bear the next 3.75% of losses, with $1.033 billion borne by private investors and $55 million by FNMA; and any further losses are retained by FNMA. With long-term, historical losses rarely exceeding 4.00%, this structure, which resembles that of most other CRTs, is designed to protect FNMA against losses other than those from its first-loss class and its skin-in-the-game vertical slices.

Finally, circle back to the relative ratings and spreads of the various CRT classes in Table 15.8. The more junior the class, the more exposed the principal to loss, the lower the rating, and the higher the spread. CRTs are bought mostly by money managers and hedge funds, with money managers favoring the more senior and hedge funds the more junior classes. Across some recent issues, about 80% to 95% of M1 classes were bought by money managers, and between about 60% and 75% of B2 classes were bought by hedge funds.[20]

The pandemic and economic shutdowns starting in 2020, bringing heightened concerns about homeowner defaults and the effects of forbearance programs, have challenged the CRT market. Both GSEs, in fact, stopped issuance for a time, but FHLMC re-entered the market in summer 2020, as did FNMA in the last quarter of 2021.

[20]Freddie Mac (2021), "Credit Risk Transfer Handbook," October, p. 23.

CHAPTER 16

Fixed Income Options

This chapter describes some of the more common fixed income options, namely: callable bonds, which have *embedded* options; Euribor futures options, which are *futures-style options*; bond futures options, which are *equity-style options*; caps and floors; and swaptions. The reader is assumed to have some familiarity with the basics of options and options pricing. Also, with the transition to LIBOR replacements still in process at the time of this writing, caps and floors and swaptions are presented in terms of LIBOR.

While fixed income options are sometimes priced using term structure models, along the lines of Chapters 7 through 9, this chapter focuses on pricing with variations of the Black-Scholes-Merton (BSM) option pricing model. This simpler approach is often used by practitioners when the objective is not to determine which options among many are relatively cheap or rich, but rather to interpolate the prices of some options from the prices of similar, but more liquid options, and to calculate usefully accurate deltas or hedge ratios. An extensive appendix justifies the use of BSM in each case with mathematically simplified versions of modern asset pricing techniques.

16.1 EMBEDDED BOND CALL OPTIONS

There is not much of a market for stand-alone options on bonds, but many corporate bonds include call provisions that give the issuer the right to call or buy back its bonds at some fixed schedule of prices. These provisions are called *embedded* options, because they do not trade separately or independently of the bonds. In any case, the relevant institutional background is given in Chapter 14. This section, for purposes of discussion, takes as an example the Bank of America (BoA) 2.305s of 02/22/2039, which were issued in February 2019. They are callable by BoA on February 22, 2029, at par, or 100% of face value. Table 16.1 lists this bond, along with two others that are used in this section as reference bonds. All three bonds are denominated in euros, and they were issued by banks that, as of the pricing date, are rated A2 by Moody's. Only the BoA bonds, however, are callable.

TABLE 16.1 Callable Bank of America Bond and Two Reference Noncallable Bonds. All Cash Flows Are Denominated in Euros. Prices Are as of August 23, 2021. Coupons Are in Percent.

Issuer	Coupon	Maturity	Call Provisions	Price
Bank of America	2.305	02/22/2039	Callable on 02/22/29 @100	113.296
DZ Bank	0.75	02/22/2029	Noncallable	104.566
Banco Santander	2.28	02/28/2039	Noncallable	118.044

The DZ Bank bonds mature on the date the BoA bonds are callable, and the Banco Santander bonds mature on the same date as the BoA bonds.

BoA will exercise its call option if rates on the call date are low relative to BoA's cost of funds. For example, if BoA can issue new 10-year debt at 1.75% as of the call date, it will likely call the 2.305s of 02/22/2039 at a price of 100 and raise the funds to do so by selling new, 10-year debt at 1.75%. In this way, BoA reduces its interest cost over the subsequent 10 years from 2.305% to 1.75%. On the other hand, if BoA's cost of new 10-year debt is 2.50% on the call date, it will not exercise its option, and it will leave outstanding its 2.305s of 02/22/2039. Return then to Table 16.1. If rates are low as of the pricing date, it is more likely that the BoA bonds will be called in 2029, which is the maturity of the DZ Bank bonds. If, on the other hand, rates are high as of the pricing date, it is less likely that the BoA bonds will be called in 2029, that is, more likely that they will remain outstanding until 2039, which is the maturity of the Banco Santander bonds. Hence, the effective maturity and duration of the BoA bonds changes from that of a 7.5-year bond, when rates are low and the call is very likely to be exercised, to that of a 17.5-year bond, when rates are high and the call is very unlikely to be exercised.

Because the BoA embedded call is a European-style option, it is particularly suitable for pricing by BSM.[1] More specifically, consider a fictional, noncallable BoA bond with a coupon of 2.305% and a maturity of February 22, 2039. The embedded call provision is then a call option on that fictional, underlying bond struck at par, and the value of the BoA callable bond equals the value of the underlying bond minus the value of the option. Essentially, bondholders own the underlying noncallable bond and have sold a call option to the issuer to buy that bond at par.

The BSM approach to pricing the BoA embedded call option is laid out in Table 16.2. Under the assumption that the forward underlying bond price

[1]A European option is exercisable at expiration. An American option is exercisable at any time between a first exercise date and expiration. And a Bermudan option is exercisable on a discrete set of dates.

TABLE 16.2 Pricing the Embedded Call Option of the Bank of America 2.305s of 02/22/2039, as of August 23, 2021.

Quantity	Value
S_0	103.732
T	7.496
K	100
σ	5.238%
$d_0(T)$	0.990
$\xi^{LN}(S_0, T, K, \sigma)$	7.877
$V_0^{BondCall} = d_0(T)\xi^{LN}(S_0, T, K, \sigma)$	7.797
Forward Rate: August 2021 to Feb 2029	0.137%
Forward Rate: Feb 2029 to Feb 2039	2.050%
Spread	−0.158%
(Fictional) Noncallable Price	122.347
Callable Price	114.551

is lognormally distributed, Appendix A16.3 shows that the value of a call on a bond is,

$$d_0(T)\xi^{LN}(S_0, T, K, \sigma) \tag{16.1}$$

where $d_0(T)$ is the discount factor to the option expiration date, T; S_0 is the forward price of the underlying bond for delivery on date T; K is the strike price; and σ is the volatility, which, under the lognormal assumption, is expressed as a percentage of price. The function ξ^{LN} is as given in Appendix A16.4.

Two of the required inputs of (16.1) are not easily available, namely the forward price to the call date of the fictional, underlying bond, and the volatility. There are several ways to estimate these missing values, but Table 16.2 takes the following approach. First, find the two forward rates – one from the pricing date to the BoA bond's call date, and one from the call date to the BoA bond's maturity date – that recover the prices of the DZ Bank and Banco Santander reference bonds. Then, using these rates, calculate that the forward price of the underlying BoA bond is 102.29. Second, assume that the relevant volatility is 59.8 basis points, which is that of an at-the-money (ATM) EUR swaption with an option expiration of seven years and an underlying tenor of 10 years, which terms are close to that of the embedded call option. Then, convert this basis-point volatility to a percentage volatility. The DV01 of the underlying forward bond, using the term structure just described, is 0.0896. Therefore, a volatility of 59.8 basis points corresponds to a price volatility of 59.8 times 0.0896, or 5.36, which, in turn, corresponds to a percentage price volatility of 5.36/102.29, or 5.24%. Not surprisingly, pricing the callable bond with this forward price and volatility does not exactly match the observed market price of 114.551 (including accrued interest). The third and final step, therefore, is to add a spread to the constructed forward rates that recovers the market price of the BoA callable bond. More precisely, for any spread, calculate the price of the underlying bond; calculate the value of the call option; and calculate the value of the callable bond as the difference between those two values. Iterate on the spread until the callable bond value equals its market price. Through this iteration, by the way, the percentage price volatility is kept constant at the 5.24% calculated previously. In any case, Table 16.2 reports the final numbers. The resulting spread is about minus 16 basis points, which increases the previously calculated underlying forward price to 103.732, and the resulting option value is 7.797. Within this framework, of course, a trader could calculate the DV01 of the callable bond, which can be used quantify, manage, or hedge its interest rate risk.

Figure 16.1 illustrates the interest rate behavior of callable bonds by graphing the prices of the three bonds in Table 16.7 against the yield of the Banco Santander bonds. The start of the pandemic and economic shutdowns is omitted, leaving pre-pandemic data, from April 30, 2019, to February 28, 2020, and post-pandemic data from May 4, 2020, to September 17, 2021. The Banco Santander and DZ Bank bonds, which are not callable, have price–yield relationships that are close to linear. These curves are

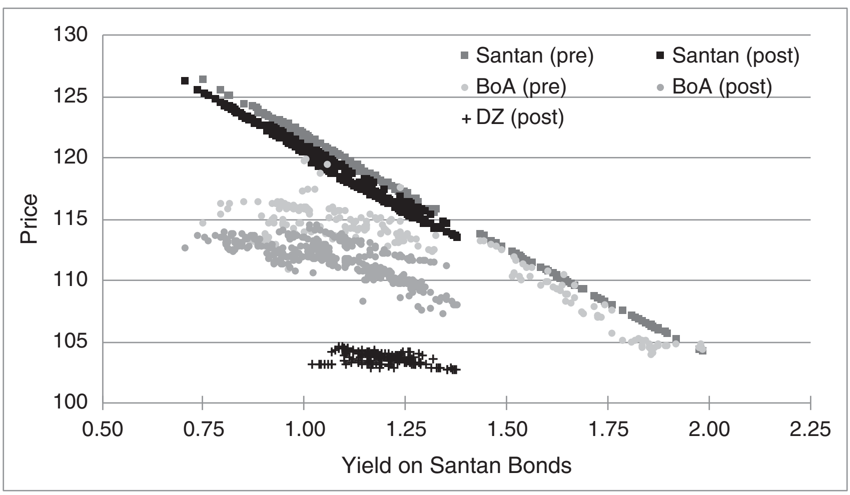

FIGURE 16.1 Prices of the Banco Santander 2.28s of 02/28/2039, the Bank of America 2.305s of 02/22/2039, and the DZ Bank 0.75s of 02/22/2029, Against the Yield of the Santander Bonds. The “Pre-Pandemic” Period Is April 30, 2019, to February 28, 2020, and the “Post-Pandemic Period” Is May 4, 2020, to September 17, 2021. The DZ Bank Bonds Were Issued in February 2021.

actually positively convex, of course, like all fixed-coupon bonds, but the nonlinearities are too small to be seen in the figure. Prices of the DZ Bank bond are relatively low, because of its low coupon. The price–yield curves of both the DZ Bank bond and the Banco Santander bond are downward sloping, of course, but the DZ curve is relatively flat, while the Santander curve is relatively steep, because the former matures in 7.5 years and the latter in 17.5 years. The price behavior of the callable BoA bond, however, is quite different, as it displays sharp and noticeable negative convexity. At high yield levels, where the likelihood of its being called in the future is low, its price–yield behavior resembles that of the similar-maturity Banco Santander bond. On the other hand, at low yield levels, where the likelihood of its being called in the future is high, its price sensitivity to rates declines: investors are likely to enjoy the now relatively high-coupon of the bond for only another 7.5 years, not for another 17.5 years. Hence, the slope of the BoA bond's price-yield curve flattens and approaches that of the 7.5-year DZ Bank bond. The BoA bonds seem to have experienced a particularly large price drop relative to Banco Santander bonds after the pandemic. This might be a divergence of issuer-specific spreads, or it might be that increased rate volatility increased the value of the embedded call and reduced the price of the callable bond.

Many callable bonds have more complex call provisions that that of the BoA bonds. For example, a new 30-year bond with a coupon of 5% might be callable in 10 years at a price of 102.50, and callable thereafter at prices that decline gradually each year until reaching 100. When this bond is issued, and the first call date is 10 years distant, most of the value of the call is at that first date and the BSM approach described in this section might be perfectly adequate. To elaborate, if the bond is unlikely to be called in 10 years at a price of 102.50, it is also unlikely to be called in 11 years at a price of, say, 102.375. Closer to the period over which the call provision is live, however, its American or Bermudan features can become more important, and the European-style assumption of BSM becomes less suitable. In these situations, it is more appropriate to apply a term structure model along the lines of Chapters 7 through 9:

1. Create an appropriate risk-neutral tree of short-term rates. In some cases the short-term rate process is designed to price bonds of the issuer selling the callable bond, but in most cases the process would correspond to a swap or government bond benchmarks. A firm- or bond-specific spread or option-adjusted spread (OAS) would then be added to rates for bond valuations.
2. Using the risk-neutral tree methodology, calculate the value along the tree of an otherwise identical, noncallable bond, that is, a noncallable bond with the same scheduled coupon and principal payments as the callable bond being priced.

3. Calculate the value along the tree of the call option embedded in the callable bond. Consistent with the well-known methods for pricing American- or Bermudan-style options along a tree, start from the maturity date of the option and work backwards, using the rules that i) the value of the option at any node equals the maximum of the value of immediately exercising the option and the value of holding the option for another period; and ii) the value of holding the option for another period at any node is its expected discounted value in the risk-neutral tree.
4. The value of the callable bond equals the value of the otherwise identical noncallable bond minus the value of the option.

With this valuation procedure in place, other computations are straightforward. Given a market price for the callable bond, an OAS can be computed along the lines of Chapter 7. Also, an *option-adjusted duration (OAD)* or DV01 can be calculated by perturbing the short-term rate factor and repeating steps 1 to 4 or, for a metric more similar to yield-based sensitivities, by perturbing the initial term structure and repeating the valuation procedure.

16.2 EURIBOR FUTURES OPTIONS

Euribor futures contracts are discussed in Chapter 12. Options on Euribor futures are *futures-style options*, because, like futures, all of their profits or losses are in the form of daily settlement payments. To explain the workings of this kind of option, consider a futures-style call option with a strike of 97.0 on some underlying futures contract. If the futures price at the expiration of the option is 99.0, then the final settlement price of the call option is 99.0 minus 97.0, or 2.0. If the futures price at expiration of the option is 95.0, then the final settlement price of the call option is zero. In other words, the final settlement price of the call option is its *intrinsic value*, that is, the maximum of zero and the difference between the underlying price and the strike.[2]

[2]The final option settlement price is actually determined through the following mechanics. Exercising a call option results in having to pay the final settlement price of the option in exchange for a position in the underlying futures contract – priced at 99 in the example – at an assigned price equal to the strike of 97. That futures position and assignment results in an immediate settlement payment of 2.0. By arbitrage, then, the final settlement price of the option must be 2.0. Through these mechanics, the option is physically settled with a position in the underlying futures contract that can be kept or sold.

Now say that a trader buys the 97 call sometime before expiration at a price of 0.20. Because this is a future-style option, the trader does not pay that 0.20 as a premium. Instead, the trader enters into the option contract like any other futures contract: the current contract price is 0.20, and all subsequent changes in the option price trigger daily settlement payments. If, at expiration, the underlying futures price is 99, the final settlement of the 97 call is 2.0, and the total of the trader's daily settlement payments will have been 2.0 minus 0.20, or 1.80. Hence, the total profit of the trader is the same 1.80 as that of a traditional call option – the final underlying price of 99, minus the strike of 97, minus the premium of 0.20. But in the case of the futures-style option, that 1.80 is realized as a sequence of daily settlement payments. Similarly, if, at expiration, the underlying futures price is 95.0, the final settlement price of the 97 call is zero, and the total of the trader's daily settlement payments will have been negative 0.20. Again, the total payoff is the same as that of a traditional call option that finishes out-of-the-money, that is, the loss of the premium, but this loss is realized in the form of daily settlement payments.

As explained in Chapter 12, the price of a Euribor futures contract is set to 100 minus its percentage rate. In the example of the previous paragraph, the strike price of 97 corresponds to a rate of (100 – 97)/100, or 3%, and the final underlying futures settlement price of 99 corresponds to a rate of 1%. Furthermore, as also explained in Chapter 12, Euribor futures contracts are scaled so that a decline of one basis point in rate is worth €25. Hence, the difference between the strike and final settlement rates, which is 3% minus 1% or 200 basis points, is worth 200 times €25, or €5,000. More generally, if the call option strike in terms of rate is K, and the final settlement rate of the underlying futures contract is R, then the final settlement price of the call option is,

$$[K - R]^+ \times 10{,}000 \times €25 = €250{,}000 \times [K - R]^+ \tag{16.2}$$

where $[K - R]^+$ is shorthand for the maximum of $K - R$ and 0, and the factor of 10,000 simply converts the rate gain, if any, to basis points, so as to be multiplied by €25. Analogously, the payoff of a put option on a rates futures contract is,

$$[R - K]^+ \times 10{,}000 \times €25 = €250{,}000 \times [R - K]^+ \tag{16.3}$$

Equations (16.2) and (16.3) show that a call option on the futures price can be expressed as a put option on its rate, and a put option on the price can be expressed as a call option on its rate. With this insight and the assumption that futures rates are normally distributed, Appendix A16.3 shows that the

TABLE 16.3 A 100.25 Call on September 2022 Three-Month Euribor Futures, as of February 28, 2022. The Futures Price Is 100.235. The Contract and Option Mature on September 19, 2022.

Quantity	Value
S_0	−0.2350%
T	$\frac{203}{365} = 0.5562$
K	−0.25%
σ	0.6842%
$\pi^N(S_0, T, K, \sigma)$	0.1961%
$V_0 = €250{,}000 \times \pi^N$	€490.34

prices of call and put options on Euribor futures are given by the following expressions, respectively,

$$€250{,}000 \times \pi^N(S_0, T, K, \sigma) \tag{16.4}$$

$$€250{,}000 \times \xi^N(S_0, T, K, \sigma) \tag{16.5}$$

where S_0 is the current futures rate; T is the time to option expiration in years; K is the strike, in terms of rate; and σ is the basis-point volatility of the futures rates. The functions π^N and ξ^N are as given in Appendix A16.4. Note that these formulae are not preceded by a discount factor, because these options are themselves futures: their prices equal their expected value under the risk-neutral measure, or equivalently, there is no initial premium on which a return need be earned. Also, the BSM model is readily applied here, because American futures-style options are not optimally exercised early: with daily settlement of the option value, there is no intrinsic value to be realized by exercise and, therefore, no reason to sacrifice the remaining time value of the option.

Table 16.3 applies Equation (16.4) to a call on the September 2022 Euribor futures contract with a strike of 100.125. As of the pricing date, the futures price is 100.235, which corresponds to a rate of −0.2350%. The futures contract and the option both mature on September 19, 2022, which is 203 days from the pricing date.

16.3 BOND FUTURES OPTIONS

Bond futures in the United States are discussed in Chapter 11. Options on these futures contracts are *traditional* or *equity-style* options, which means that, like options on stocks, purchasers of options pay a premium at the time of purchase and, if exercised, realize their intrinsic values.

Options on bond futures can be valued with term structure models. In a one-factor model, for example, after creating a risk-neutral tree for the futures price, the value of a futures option can be computed by starting from the maturity date of the option and working backwards, using the rule that the value of the option at each node is the maximum of the value of holding the option and of exercising it immediately. Of course, because the delivery options embedded in bond futures may not be modeled adequately by a one-factor model, multi-factor approaches might be preferred.

While simple conceptually, it is apparent from Chapter 11 that building a futures model takes a good deal of effort. Therefore, practitioners who do not otherwise need such a model tend to use a BSM approach for bond futures options. The required assumptions are that: the option is European; the discount factor to option expiration is uncorrelated with the underlying futures price; and the futures price is lognormal with constant volatility. The assumption that the option is European, that is, that early exercise is not optimal, is reasonable in practice. Early exercise of an equity-style futures option may be optimal if interest earned on the realized intrinsic value exceeds the time value of the option, but this is rarely the case. For some intuition on this point, note that early exercise of a put on a stock may be optimal, because of the potential of earning interest on the entire strike price, not just on its intrinsic value. The assumption that the discount factor is uncorrelated with the bond futures price is not bad in practice. Bond futures options typically mature in a few months or less, while the bond underlying the futures contract typically matures in many years. Hence, the correlation is typically weak between the discount factor, which depends on very short-term rates, and the bond futures price, which depends on longer-term rates. The assumption of constant price volatility, however, essentially assumes away the delivery options of bond futures: as rates change and the cheapest-to-deliver bond changes, the DV01 and, therefore, the price volatility of the futures changes as well. (See Chapter 11.) While perhaps acceptable when interest rates are low relative to the contract's notional coupon, ignoring the delivery option seems to undercut one of the main motivations for using BSM in the first place, namely, to obtain accurate deltas. Nevertheless, BSM is used in practice in this context.

Under the assumptions listed in the previous paragraph, Appendix A16.3 shows that calls and puts on US Treasury note and bond futures contracts are given by, respectively,

$$\$N \times d_0(T) \times \xi^{LN}(S_0, T, K, \sigma)/100 \tag{16.6}$$

$$\$N \times d_0(T) \times \pi^{LN}(S_0, T, K, \sigma)/100 \tag{16.7}$$

TABLE 16.4 A Call Option on the June 10-Year Treasury Note Futures Contract, as of February 28, 2022. The Futures Contract and Option Mature on June 21, 2022.

Quantity	Value
S_0	126.3438
T	$\frac{113}{365} = 0.3096$
K	126.5
σ	5.655%
$d_0(T)$	0.9979%
$\xi^{LN}(S_0, T, K, \sigma)$	1.50996
$V_0 = \$100{,}000 \times d_0(T) \times \xi^{LN}/100$	$1,509.96

where N is the face amount of bonds per contract; S_0 is the futures price; T is the time to option expiration; K is the strike price; σ is the lognormal or percentage price volatility; and $d_0(T)$ is the discount factor to option expiration. Table 16.4 applies Equation (16.6) to a call option on the June 10-year US Treasury note futures contract as of the end of February 2022, where both the futures and option contracts expire on June 21, 2022. Note that there are 113 days from February 28, 2022, to June 21, 2022.

16.4 CAPS AND FLOORS

The easiest way to describe *caps* is to start with *caplets*, even though caps are the more traded derivative. At the end of a given accrual period, a caplet pays the greater of zero and of LIBOR minus a strike, where LIBOR is set at the beginning of the accrual period. Consider, for example, a caplet with a three-month LIBOR reset date of February 14, 2022, a payment date of May 14, 2022, and a strike of .181%. Note that there are 89 days over this accrual period. If LIBOR on February 14, 2022, turns out to be L, then a unit notional of the caplet will pay, on May 14, 2022,

$$\frac{89}{360}[L - .181\%]^+ \tag{16.8}$$

where, as mentioned already, $[L - .181\%]^+$ is another way of writing max $[L - 0.181\%, 0]$. Note that the payoff of a caplet looks like that of an option, but the maximum is a feature of the contract rather than a result of optimal exercise behavior.

Caplets are typically valued by practitioners under the assumption that forward LIBOR rates are normally distributed. Under this assumption,

Appendix A16.3 shows that the value of a caplet with a reset at time T and payment at time $T + \tau$ is given by,

$$\tau d_0(T + \tau)\xi^N(S_0, T, K, \sigma) \tag{16.9}$$

where τ is the term of the reference rate; $d_0(T + \tau)$ is the discount factor to the payment date; S_0 is today's forward rate from T to $T + \tau$; K is the strike; σ the basis-point volatility of the forward rate; and ξ^N the BSM-style formula defined in Appendix A16.4. Table 16.5 applies Equation (16.9) to 100 notional of the caplet introduced previously, as of May 14, 2021. Note that there are 276 days from the pricing date, May 14, 2021, to the reset date, February 14, 2022.[3] The appropriate discount factor to the payment date, derived from the swap curve, is .998191. Finally, a volatility of 12.09 basis points, which comes from the discussion, is used to derive the price of .0129 cents per 100 notional amount.

A cap is a portfolio of caplets, with the value of the cap being the sum of the value of its component caplets. The implied volatility of a cap is the volatility that, when used to value every component caplet, results in the market price of the cap. This leads to some complexities, as will be discussed presently, because the term structure of caplet volatility is not flat. In other words, every caplet is properly valued at its own volatility even though, when quoting the price of a cap, all of its component caplets are valued at a single, cap volatility.

TABLE 16.5 Pricing a Caplet with a LIBOR Reset on February 14, 2022, and a Payment Date on May 14, 2022, as of May 14, 2021.

Quantity	Value
S_0	.200%
T	$\frac{276}{365} = .7562$
τ	$\frac{89}{360} = .2472$
K	.181%
σ	.1209%
$d_0(T+\tau)$	.998191
$\xi^N(S_0, T, K, \sigma)$	.00052
$V_0^{Caplet} = 100 \times \tau d_0(T+\tau)\xi^N$	.0129

[3]For simplicity, this presentation does not distinguish between business and non-business days.

Table 16.6 illustrates the structure of a cap and cap pricing with a one-year US dollar ATM cap as of May 14, 2021. This cap is ATM because its strike of 0.181% equals the rate of the corresponding swap, which, in this case, is the one-year swap. The cap strike of 0.181% means that every component caplet has a strike of 0.181%. The cap volatility of 12.09 basis points means that the price of the cap is the sum of the component caplet values when each caplet is valued at a volatility of 12.09 basis points. The forward rates are derived from the swap curve, and the caplet premiums are calculated from the normal BSM formula, with: each respective forward rate; a strike of 0.181%; a volatility of 12.09 basis points; and appropriate date parameters and discount factors along the lines of Table 16.5. In fact, the caplet in Table 16.6 that pays on May 14, 2022, is the same caplet that is valued in Table 16.5.

Note that what might have been the first caplet in Table 16.6, with a LIBOR reset at the start of the cap initiation and a payment on August 14, 2021, is omitted from the table. The payment from such a caplet is known as of the start of the cap and, as such, has no option-like premium: it is simply worth its present value. In fact, in this example, where the initial LIBOR setting is below the strike, at .155%, the payment from this caplet is zero. In any case, along these lines, the first caplet payment is usually omitted from a cap.

The one-year cap in the example is a spot starting cap; that is, putting aside the skipping of the first payment, the schedule of payments starts immediately. There is also an active market, however, in forward starting caps. In a 5×5 cap, the first reset is in five years, the first payment in five years plus the length of the accrual period (e.g., five years and three months), and the last payment in 10 years.

As mentioned before, if caplets traded individually, they would be priced at individual volatilities, not at a single, cap volatility. In other words, there

TABLE 16.6 The Structure and Pricing of a One-Year Cap as of May 14, 2021. Rates Are in Percent.

Cap Strike			**.181%**
Cap Volatility			**.1209%**
Dates			
Reset	**Payment**	**Forward Rate**	**Caplet Premium**
05/14/2021	08/14/2021	.155	
08/14/2021	11/14/2021	.160	.0039
11/14/2021	02/14/2022	.200	.0114
02/14/2022	05/14/2022	.200	.0129
Sum			.0281

is a term structure of caplet volatilities. This term structure is interesting for use as another perspective on the market price of volatility and for comparison with – and perhaps trading opportunities against – other volatility instruments. In theory, the term structure of caplet volatility could be recovered from caps of sequential terms. The problem is complicated, however, by the fact that the most traded and useful volatilities are ATM volatilities, which, in the case of caplets, correspond to caplets with strikes equal to their underlying forward rates. But the strikes of caplets that are part of caps all have a single strike that cannot, in general, equal the underlying forward rate for every component cap. Furthermore, as discussed presently, the volatilities of options that are and are not ATM can be significantly different from each other. Hence, the extraction of caplet volatilities from caps is often combined with some adjustment for the caplet strikes not being ATM.

Floorlets and floors are analogous to caplets and caps. The payment of a floorlet at time $T + \tau$, determined by the LIBOR rate set at time T, is,

$$\tau[K - L]^{+} \tag{16.10}$$

Assuming normal forward rates, the value of a floorlet is given by,

$$\tau d_0(T + \tau)\pi^N(S_0, T, K, \sigma) \tag{16.11}$$

where the function π^N is given in Appendix A16.4. The value of a floor is the sum of the values of its component floorlets.

Applying put-call parity, the prices of an ATM caplet and an ATM floorlet with the same expiration are equal, as are the prices of matched-date ATM caps and floors. Say, for example, that the five-year swap rate, five years forward, is 5%. Then paying fixed on this forward starting swap and buying a 5×5 floor with a strike of 5% has exactly the same cash flows as a 5×5 cap with a strike of 5%. But, by definition, the value of the forward swap is zero. Hence, the values of the cap and the floor must be the same.

16.5 SWAPTIONS

A *swaption* is an over-the-counter (OTC) contract that gives the buyer the right, at expiration, to enter into a fixed-for-floating interest rate swap of a prespecified maturity and strike rate. For example, a "2-year-5-year" or "2y5y" swaption is a two-year option to enter into a five-year swap at some prespecified strike. A *receiver swaption* gives the buyer the right to receive fixed and pay floating, while a *payer swaption* gives the buyer the right to pay fixed and receive floating.

For presenting the pricing of swaptions, consider a \$100 million 2.36% 5y5y receiver swaption traded on May 14, 2021. This option gives the buyer the right, in five years, on May 14, 2026, to receive 2.36% and pay LIBOR on $100 million for five years, that is, until May 14, 2031. To write down the value of this swaption at expiration, let $C_5(5,10)$ denote the five-year par swap rate, five years from today (which matures in 10 years), and let $A_5(5,10)$ denote the value five years from today of an annuity of $1 per year, paid on each of the fixed-rate payment dates of a five-year swap starting in five years (and ending in 10 years). Then, in five years, at the expiration of the swaption, the value of receiving 2.36% for five years is,

$$\$100 \text{ mm} \times [2.36\% - C_5(5,10)]^+ \times A_5(5,10) \tag{16.12}$$

Inspection of the payoff (16.12) reveals that a 5y5y receiver swaption is a put on the five-year par swap rate, five year forward. More generally, a T-year-τ-year receiver swaption is a T-year put option on the τ-year par swap rate, T-years forward. Similarly, a T-year-τ-year payer swaption is a T-year call option on the τ-year par swap rate, T-years forward.

Table 16.7 applies BSM to the example of this section. As just discussed, the rate underlying the 2.36% 5y5y receiver option traded on May 14, 2021, is the forward par rate on a swap from May 14, 2026, to May 14, 2031. Hence S_0 of BSM is 2.36%, and the swaption of the example, with its strike at 2.36%, is ATM.[4] For the 2.36% 5y5y, the other parameters are clearly $T = 5$, $\tau = 5$, and $K = 2.36\%$. The value of the annuity on the swap from May 14, 2026, to May 14, 2031, as of May 14, 2021, is 4.287. Finally, the market price of this receiver option on the pricing date is 3.03 per 100 notional amount or $3.03 million on $100 million. Appendix A16.3 shows that, when the underlying forward swap rate is normally distributed, the value of a receiver swaption per unit notional is,

$$A_0(T, T+\tau) \times \pi^N(S_0, T, K, \sigma) \tag{16.13}$$

where π^N is once again from Appendix A16.4. Setting the market price of $3.03 million equal to $100 million times (16.13), Table 16.7 shows that the implied volatility of this swaption is 0.793%.

The analogous formula for a payer swaption per unit notional is,

$$A_0(T, T+\tau) \times \xi^N(S_0, T, K, \sigma) \tag{16.14}$$

The prices of ATM swaptions, which are by far the most commonly traded swaptions, are quoted in a matrix of either premia or implied normal volatilities. Table 16.8 is an example of the latter for US dollar swaptions

[4] A higher strike would make the receiver swaption *in-the-money*, while a lower strike would make it *out-of-money*

TABLE 16.7 A 5y5y Receiver Swaption per 100 Notional Amount of Swaps, as of May 14, 2021.

Quantity	Value
S_0	2.36%
T	5
τ	5
K	2.36%
σ	0.793%
$A_0(T, T+\tau)$	4.287
$\pi^N(S_0, T, K, \sigma)$	0.007074
$V_0^{Receiver} = \$100 \text{ mm} \times A \times \pi^N$	3.03 mm

TABLE 16.8 USD ATM Swaption Normal Volatilities in Basis Points, as of May 14, 2021.

Option	Swap Tenor										
Exp.	1y	2y	3y	4y	5y	7y	10y	12y	15y	20y	30y
1m	12.1	18.6	34.1	45.9	57.1	64.7	70.7	71.0	71.1	71.3	71.5
3m	14.1	23.3	39.8	51.7	62.5	69.2	74.4	74.5	74.4	74.3	74.3
6m	16.7	29.4	45.5	56.0	66.4	72.0	76.3	76.2	75.8	75.6	75.3
1y	29.3	44.1	57.0	64.1	70.3	73.9	76.8	76.5	75.8	75.0	74.2
2y	59.2	68.9	73.3	76.0	77.2	77.5	77.7	77.0	75.7	74.1	72.9
3y	76.7	79.6	79.5	79.4	79.3	78.6	77.6	76.7	75.1	72.8	71.2
4y	81.4	81.2	80.7	80.2	79.8	78.4	76.6	75.5	73.7	71.3	69.4
5y	82.3	81.2	80.5	79.9	79.3	77.7	75.4	74.1	72.2	69.9	67.5
10y	73.5	72.3	71.8	71.2	70.8	69.2	66.8	65.6	63.8	61.6	59.4

as of May 14, 2021. For example, the ATM 2y10y options is priced with an implied volatility of 77.7 basis points.[5] The price of the 5y5y option introduced previously is quoted at a volatility of 79.3 basis points.

Swaption *skew*, discussed in the next section, refers to the fact that implied volatilities for ATM swaptions in Table 16.8 are valid only for ATM swaptions. The broader swaptions market, therefore, actually trades a volatility *cube*, where the third dimension represents strike, usually in 50 basis-point increments away from the forward swap rate corresponding to each entry of the swaption matrix. For a 5y5y as of May 14, 2021, with the underlying par forward swap at 2.36%, a volatility cube would show

[5] ATM calls and puts have the same BSM values. This is easy to verify from Equations (A16.60) through (A16.66).

volatilities for the 5y5y at higher strikes of 2.86%, 3.36%, and so forth, and for lower strikes of 1.86%, 1.36%, etc.

The skew applies not only for trading swaptions that are not ATM but also for valuing or marking existing swaptions that were initiated as ATM options. Consider the 2.36% 5y5y receiver swaption traded on May 14, 2021. The underlying swap of this swaption, from initiation to expiration, is a 2.36% swap from May 14, 2025, to May 14, 2031. As of May 14, 2021, this swap is a par swap, but over time this will no longer be the case. Say, for example, that one month later, on June 14, 2021, the rate of the forward par swap corresponding to those dates is 2.50%. In that case, the 2.36% receiver swaption can be characterized as a 4-year-11-month-5-year that is 14 basis points out-of-the-money. Valuing the option, therefore, requires an interpolation between the ATM and the 50 basis-point, out-of-the-money volatilities, as well as an interpolation between the 4y5y and 5y5y option expirations. In practice, these interpolations are carried out by means of a stochastic volatility model, discussed later in this chapter.

This section describes swaptions as if they are physically settled, meaning that, at expiration, the counterparties enter into a swap at the appropriate rate and maturity. In fact, however, swaptions in the United States are almost always cash settled and, in Europe, can be physically or cash settled. In the United States, the cash settlement is found by multiplying the appropriate annuity factor, evaluated along the swap curve, by the difference between the par rate and the strike, in the case of payers, or the difference between the strike and the par rate, in the case of receivers. In Europe, the annuity factor is computed at a flat rate equal to the appropriate par swap rate, which can lead to very minor valuation differences between the two forms of settlement. As a final note, for some swaptions in Europe, the premium is paid at the expiration date rather than the trade date, which changes valuations accordingly.

16.6 SWAPTION SKEW

The BSM normal model, as applied in the previous section, has a single and constant volatility parameter. Taking the model literally, this would imply that all swaptions can be priced with a single basis-point volatility. Figure 16.2 illustrates, however, that this is not the case: even across swaptions with the same expiration and underlying tenor, like either the 2y2y and 5y10y swaptions in the figure, implied volatilities vary significantly with strike. As of the pricing date, the two-year swap rate, two years forward is 1.11% (not shown in the figure), which is the ATM strike for 2y2y swaptions. The darker gray curve shows, therefore, that 2y2y implied volatilities increase mostly linearly from 56 basis points at a strike of 0.11%, which is 100 basis points below the ATM strike, to 125 basis points at a strike

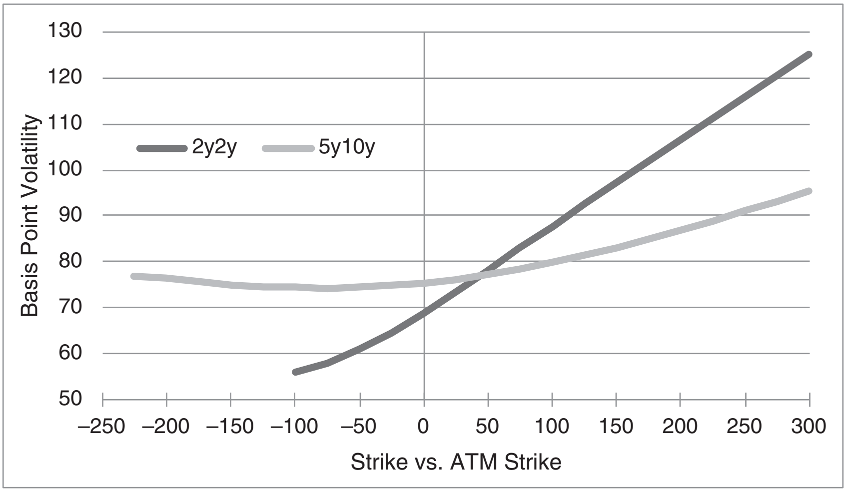

FIGURE 16.2 Implied Basis-Point Volatilities of 2y2y and 5y10y US Dollar Swaptions Across Strikes, as of May 14, 2021.

of 4.11%, which is 300 basis points above the ATM strike. At the same time, the 10-year swap rate, five years forward is 2.43%. The lighter gray curve of the figure shows, therefore, that implied volatilities of 5y10y swaptions are roughly 75 basis points for strikes below that ATM strike, and then increase along the curve shown to 95 basis points at a strike 300 basis points over the ATM strike. The phenomenon that basis-point volatility is not constant across strikes is known as the volatility *skew*. The phenomenon that basis-point volatility is higher for both in- and out-of-the-money options than for at-the-money options, as seen here for 5y10y swaptions, is known as the *smile*, or, to the extent this effect is not symmetric, the *smirk*.

The existence of a skew does not prevent traders from quoting swaption prices in terms of normal BSM volatility or *vice versa*, so long as the volatilities vary with strike. But practitioners rely on BSM not just to quote a price or volatility, but also to compute delta, that is, to compute how the value of a swaption changes as its underlying forward rate changes. And the usefulness of BSM for this purpose is cast into doubt by Figure 16.2. Without getting into great detail here, the implied volatility of an option at a given strike can be thought of as the expected gamma-weighted average of instantaneous volatilities over possible paths of the underlying forward rate as it changes from its current level to the strike. From this perspective, the implied volatility curves in Figure 16.2 mean that instantaneous volatilities vary in a complex way with the level the forward rate. In that case, however, the value of an option changes with the level of rates in two ways: a direct effect, that is, the change in the underlying forward; and an indirect effect, that is, the change in volatility as a result of the change in the level of rates.

But the BSM delta, which is traditionally computed as the change in option value for a change in the underlying at a given volatility, captures only the direct effect.

The search for a model that can capture the underlying dynamics of volatility as a function of forward rates has taken two very broad paths. The first is to change the distribution of the forward rate away from normal. For example, a constant-volatility lognormal model of rates assumes that volatility is proportional to rates, which might be useful in explaining some manifestations of the skew, like parts of the 2y2y curve in Figure 16.2. In any case, along these lines, the *shifted lognormal model* allows for the distribution of the underlying forward rate to be between normal and lognormal. The defining dynamics of the forward rate, S, in this model are,

$$dS_t = \phi(S_t)\sigma\epsilon_t\sqrt{dt} \tag{16.15}$$

$$\epsilon_t \sim N(0,1) \tag{16.16}$$

$$\phi(S_t) = a + S_t \tag{16.17}$$

with $a \geq 0$. When $a = 0$, $\phi(S_t) = S_t$, the basis-point volatility in Equation (16.15) is just σS_t, that is, a constant proportion of the forward rate. Therefore, the distribution of the rate is lognormal. At the other extreme, as a approaches infinity, $\phi(S_t)$ in (16.17) approaches the constant a, which means that the volatility in (16.15) approaches a constant and, therefore, that the distribution approaches normality.

The second approach to finding a model that captures the relationship between rates and volatility is to allow volatility itself to be a random variable. This approach leads to *stochastic volatility* models, in which class the *SABR* model had proven particularly popular. As originally formulated, however, the model could not handle the negative rates that recently characterized markets in Europe. An adjusted model then emerged, known as the *shifted-SABR* model, which assumes the following dynamics,

$$dS_t = (S_t + h)^\beta \sigma_t \epsilon_t \sqrt{dt} \tag{16.18}$$

$$\epsilon_t \sim N(0,1) \tag{16.19}$$

$$d\sigma_t = \alpha\sigma_t \nu_t \sqrt{dt} \tag{16.20}$$

$$\nu_t \sim N(0,1) \tag{16.21}$$

$$E[\epsilon_t \nu_t] = \rho \tag{16.22}$$

with $h \geq 0$, $0 < \beta < 1$, $\alpha \geq 0$, and $0 \leq \rho \leq 1$. There are several features to note about this formulation. First, the only difference between this model and the original SABR model is the shift parameter, h. With this shift, the forward rate S_t can be as negative as $-h$ with the basis-point volatility of the model – $(S_t + h)^\beta \sigma_t$ – still positive. Second, the SABR model approaches the normal model as both α and β approach 0, and it approaches the lognormal model as α approaches 0 and as β approaches 1. Third, the initial volatility, σ_0, is most naturally used to match ATM swaption volatility. Fourth, β is typically less than one, because basis-point volatilities, as illustrated in Figures 16.2, do not increase as quickly for high strikes as in a lognormal model. Fifth, the parameters β and ρ control the skew through controlling the relationship between the level of rates and volatility. From Equation (16.18), a higher β increases the responsiveness of volatility to the level of the forward rate. And from Equation (16.22), a higher ρ increases the correlation between changes in rates and changes in volatility. Sixth, the parameter α controls the smile, or *fat tails* of the rates distribution, as a higher α increases volatility when volatility is high, that is, α increases volatility when rates are either higher or lower. Seventh, in practice, because the flexibility of the model enables many different sets of parameters to fit the observed skew, the parameters can also be constrained to fit empirical regularities. One popular choice, for example, is to constrain the parameters to match the empirical *backbone*, that is, the empirical relationship between ATM basis-point volatilities and forward rates. Eighth, note that, in practice, a different set of parameters are typically chosen for every swaption expiration and underlying tenor. In other words, no BSM-style model can describe the entire volatility cube.

APPENDIX TO CHAPTER 1

Prices, Discount Factors, and Arbitrage

A1.1 DERIVING REPLICATING PORTFOLIOS

To replicate the 7.625s of 11/15/2022, Table 1.5 uses the 2.875s of 11/15/2021, the 2.125s of 05/15/2022, and the 1.625s of 11/15/2022. Number these bonds from 1 to 3, and let F_i be the face amount of bond i in the replicating portfolio. Then, the following equations express the requirement that the cash flows of the replicating portfolio equal those of the 7.625s on each of the cash flow dates. For the cash flows on November 15, 2021,

$$\left(100\% + \frac{2.875\%}{2}\right) F_1 + \frac{2.125\%}{2} F_2 + \frac{1.625\%}{2} F_3 = \frac{7.625\%}{2} \tag{A1.1}$$

For the cash flows on May, 15, 2022,

$$0 \times F_1 + \left(100\% + \frac{2.125\%}{2}\right) F_2 + \frac{1.625\%}{2} F_3 = \frac{7.625\%}{2} \tag{A1.2}$$

And for the cash flows on November 15, 2022,

$$0 \times F_1 + 0 \times F_2 + \left(100\% + \frac{1.625\%}{2}\right) F_3 = 100\% + \frac{7.625\%}{2} \tag{A1.3}$$

Solving equations (A1.1), (A1.2), and (A1.3) for F_1, F_2, and F_3 gives the replicating portfolio's face amounts reported in Table 1.5. Note that, because one bond matures on each date, these equations can be solved one-at-a-time instead of simultaneously.

Replicating portfolios are easier to describe and manipulate using matrix algebra. To illustrate, equations (A1.1) through (A1.3) can be written as follows,

$$\begin{bmatrix} 1+\frac{2.875\%}{2} & \frac{2.125\%}{2} & \frac{1.625\%}{2} \\ 0 & 1+\frac{2.125\%}{2} & \frac{1.625\%}{2} \\ 0 & 0 & 1+\frac{1.625\%}{2} \end{bmatrix} \begin{bmatrix} F_1 \\ F_2 \\ F_3 \end{bmatrix} = \begin{bmatrix} \frac{7.625\%}{2} \\ \frac{7.625\%}{2} \\ 1+\frac{7.625\%}{2} \end{bmatrix} \tag{A1.4}$$

Note that each column of the leftmost matrix describes the cash flows of one of the bonds in the replicating portfolio; the elements of the vector to the right of the matrix gives the face amounts of each bond for which the equation has to be solved; and the rightmost vector gives the cash flows of the bond to be replicated. Equation (A1.4) can easily be solved.

In general, suppose that the bond to be replicated makes payments on T dates. Let **C** be the $T \times T$ matrix of cash flows, principal plus interest, with the T columns representing the T bonds in the replicating portfolio and the T rows the dates on which those bonds make payments. Let **f** be the $T \times 1$ vector of face amounts in the replicating portfolio, and let **c** be the vector of cash flows, principal plus interest, of the bond to be replicated. Then, the equation to be solved is,

$$\mathbf{Cf} = \mathbf{c} \tag{A1.5}$$

with solution,

$$\mathbf{f} = \mathbf{C}^{-1}\mathbf{c} \tag{A1.6}$$

The only complication in constructing replicating portfolios is to ensure that the matrix **C** does have an inverse. Essentially, any set of T bonds will do so long as there is at least one bond in the replicating portfolio making a payment on each of the T dates. All T bonds maturing on the last date would work, for example, but all T bonds maturing on the second-to-last date would not. In the latter case, there would be no bond in the replicating portfolio making a payment on date T.

A1.2 THE EQUIVALENCE OF DISCOUNTING AND ARBITRAGE PRICING

Proposition: Pricing a bond according to either of the following methods gives the same price:

1. Derive a set of discount factors from some set of spanning bonds and price the bond in question using those discount factors.

2. Find the replicating portfolio of the bond in question using that same set of spanning bonds and calculate the price of the bond as the price of this portfolio.

Proof: Continue with the notation introduced at the end of Appendix A1.1. In addition, let $\mathbf{d}$ be the $T \times 1$ vector of discount factors for each date and let $\mathbf{p}$ be the vector of prices of each bond in the replicating portfolio, which is the same as the vector of prices used to compute the discount factors. Also note that a set of spanning bonds is such that the matrix $\mathbf{C}$ has an inverse.

Generalizing the derivation of discount factors in this chapter, discount factors can be determined from the following equation,

$$\mathbf{d} = \mathbf{C}'^{-1}\mathbf{p} \tag{A1.7}$$

where the $'$ denotes the transpose. Then, the price of the bond according to the first method is $\mathbf{c}'\mathbf{d}$. The price according to the second method is $\mathbf{p}'\mathbf{f}$, where $\mathbf{f}$ is as derived in Equation (A1.6).

Hence, the two methods give the same price if,

$$\mathbf{c}'\mathbf{d} = \mathbf{p}'\mathbf{f} \tag{A1.8}$$

Expanding the left-hand side of Equation (A1.8) with (A1.7) and the right-hand side with (A1.6),

$$\mathbf{c}'(\mathbf{C}')^{-1}\mathbf{p} = \mathbf{p}'\mathbf{C}^{-1}\mathbf{c} \tag{A1.9}$$

Finally, since both sides of (A1.9) are numbers, take the transpose of the left-hand side to show that the equation is true.

APPENDIX TO CHAPTER 2

Swap, Spot, and Forward Rates

A2.1 CONTINUOUS COMPOUNDING

Equation (2.7) gives the proceeds of investing F for N periods at the rate $\hat{r}$, which is compounded n times per year. By the definition of n, there are nT periods over T years. Therefore, with $F = 1$, (2.7) becomes,

$$\left(1 + \frac{\hat{r}}{n}\right)^{nT} \tag{A2.1}$$

Under continuous compounding, interest is paid every instant, so that the proceeds of an investment that is continuously compounded over T years grows to the limit of Equation (A2.1) as n approaches infinity. Taking the logarithm of (A2.1) and rearranging terms,

$$nT \ln\left(1 + \frac{\hat{r}}{n}\right) = \frac{T \ln(1 + \frac{\hat{r}}{n})}{\frac{1}{n}} \tag{A2.2}$$

Using l'Hôpital's rule, the limit of the right-hand side of (A2.2) as n becomes large is $\hat{r}T$. Hence, the limit of (A2.1) is $e^{\hat{r}T}$, where e is the base of the natural logarithm. Therefore, if interest is continuously compounded at the rate $\hat{r}$, an investment of one unit of currency will grow after T years to,

$$e^{\hat{r}T} \tag{A2.3}$$

Equivalently, the present value of one unit of currency to be received in T years is,

$$e^{-\hat{r}T} \tag{A2.4}$$

This section now defines discount factors, spot rates, and forward rates under continuous compounding. Let $\hat{r}(t)$ be the t-year continuously compounded spot rate. Let $f(t - \Delta, t)$ be the forward rate from $t - \Delta$ to t, and define $f(t)$ to be the continuously compounded forward rate at time t, that is, the limit of $f(t-\Delta,t)$ as Δ approaches zero.

By Equation (A2.4), spot rates and discount factors are related such that,

$$d(T) = e^{-\hat{r}(T)T} \tag{A2.5}$$

Linking forward rates and spot rates is the continuously compounded analog of Equation (2.21),

$$e^{\hat{r}(T)T} = e^{\int_0^T f(s)ds} \tag{A2.6}$$

$$\hat{r}(T)T = \int_0^T f(s)ds \tag{A2.7}$$

$$\hat{r}(T) = \frac{1}{T}\int_0^T f(s)ds \tag{A2.8}$$

To link forward rates and discount factors, note that the continuously compounded analogue of Equation (2.20) is,

$$e^{\hat{r}(t-\Delta)\times(t-\Delta)}e^{f(t-\Delta,t)\Delta} = e^{\hat{r}(t)t} \tag{A2.9}$$

Then substitute from Equation (A2.5) for each of the two spot rates and rearrange terms to get,

$$e^{f(t-\Delta,t)\times\Delta} = \frac{d(t-\Delta)}{d(t)} \tag{A2.10}$$

Then take the natural logarithm of both sides and rearrange terms,

$$f(t-\Delta,t) = -\frac{\ln[d(t)] - \ln[d(t-\Delta)]}{\Delta} \tag{A2.11}$$

Finally, take the limit of both sides of this equation, recognizing that the limit of the right-hand side is the derivative of $\ln[d(t)]$, to obtain,

$$f(t) = -\frac{d'(t)}{d(t)} \tag{A2.12}$$

where $d'(t)$ is the derivative of the discount function with respect to term.

A2.2 RELATIONSHIPS BETWEEN SWAP OR PAR, SPOT, AND FORWARD RATES

This section will work with semiannually compounded rates, though it could easily be cast in terms of other compounding intervals.

Approximation: The t-year spot rate is approximately equal to the average of all forward rates to year t.

Start from Equation (2.21), noting that, because interest rates themselves are small numbers, the product of two or more interest rates is particularly small. To illustrate, take the case of the one-year spot rate, though the argument generalizes easily to longer-term rates,

$$\left(1+\frac{\hat{r}(1.0)}{2}\right)^2=\left(1+\frac{f(0.5)}{2}\right)\left(1+\frac{f(1)}{2}\right) \tag{A2.13}$$

$$1+2\frac{\hat{r}(1.0)}{2}+\left(\frac{\hat{r}(1.0)}{2}\right)^2=1+\frac{f(0.5)}{2}+\frac{f(1.0)}{2}+\frac{f(0.5)}{2}\frac{f(1.0)}{2} \tag{A2.14}$$

$$\hat{r}(1.0)\approx\frac{f(0.5)+f(1.0)}{2} \tag{A2.15}$$

where the approximation from (A2.14) to (A2.15) comes from dropping the terms that multiply two rates.

Proposition 1:

$$\sum_{t=a}^{b} z^t=\frac{z^a-z^{b+1}}{1-z} \tag{A2.16}$$

Proof: Define S as the left-hand side of (A2.16). Then,

$$zS=\sum_{t=a+1}^{b+1} z^t$$

and it follows that,

$$S-zS=\sum_{t=a}^{b} z^t-\sum_{t=a+1}^{b+1} z^t$$

$$S(1-z)=z^a-z^{b+1}$$

$$S=\frac{z^a-z^{b+1}}{1-z}$$

as was to be shown.

Proposition 2: If the term structure of spot rates is flat, then the term structure of par rates is flat at that same rate.

Proof: Denote the semiannually compounded par rate of maturity T as $C(T)$. If spot rates are flat at the rate $\hat{r}$, then, by definition of $C(T)$,

$$\frac{C(T)}{2}\sum_{t=1}^{2T}\frac{1}{\left(1+\frac{\hat{r}}{2}\right)^{t}}+\frac{1}{\left(1+\frac{\hat{r}}{2}\right)^{2T}}=1 \tag{A2.17}$$

Applying Equation (A2.16) of Proposition 1 with $z = 1/(1+\hat{r}/2)$,

$$\frac{C(T)}{\hat{r}}\left[1-\frac{1}{(1+\frac{\hat{r}}{2})^{2T}}\right]+\frac{1}{\left(1+\frac{\hat{r}}{2}\right)^{2T}}=1 \tag{A2.18}$$

But solving (A2.18) for $C(T)$ shows that $C(T) = \hat{r}$. Hence, the term structure of par rates is flat at $\hat{r}$, as was to be shown.

Proposition 3: $f(t) > \hat{r}(t-0.5)$ if and only if $\hat{r}(t) > \hat{r}(t-0.5)$.

Proof: The condition $f(t) > \hat{r}(t-0.5)$ is equivalent to,

$$\left(1+\frac{\hat{r}(t-0.5)}{2}\right)^{2t-1}\left(1+\frac{f(t)}{2}\right) > \left(1+\frac{\hat{r}(t-0.5)}{2}\right)^{2t-1}\left(1+\frac{\hat{r}(t-0.5)}{2}\right)$$
$$> \left(1+\frac{\hat{r}(t-0.5)}{2}\right)^{2t} \tag{A2.19}$$

But, using Equation (2.20) to rewrite the left-hand side of (A2.19),

$$\left(1+\frac{\hat{r}(t)}{2}\right)^{2t} > \left(1+\frac{\hat{r}(t-0.5)}{2}\right)^{2t} \tag{A2.20}$$
$$\hat{r}(t) > \hat{r}(t-0.5) \tag{A2.21}$$

as was to be shown.

Proposition 4: $f(t) < \hat{r}(t-0.5)$ if and only if $\hat{r}(t) < \hat{r}(t-0.5)$.

Proof: Reverse the inequalities in the proof of Proposition 3.

Proposition 5: If $\hat{r}(0.5) < \hat{r}(1.0) < \cdots < \hat{r}(T)$, then $C(T) < \hat{r}(T)$.

Proof: By the definition of the par rate, $C(T)$,

$$\frac{C(T)}{2}\sum_{t=1}^{2T}\frac{1}{\left(1+\frac{\hat{r}(t)}{2}\right)^{t}}+\frac{1}{\left(1+\frac{\hat{r}(T)}{2}\right)^{2T}}=1 \tag{A2.22}$$

It is easy to show from Equation (A2.16), setting $z = 1/(1 + C(T)/2)$, that,

$$\frac{C(T)}{2}\sum_{t=1}^{2T}\frac{1}{\left(1+\frac{C(T)}{2}\right)^{t}}+\frac{1}{\left(1+\frac{C(T)}{2}\right)^{2T}}=1 \tag{A2.23}$$

And, because the term structure of spot rates is assumed to be increasing,

$$\begin{aligned}\frac{C(T)}{2}\sum_{t=1}^{2T}\frac{1}{(1+\frac{\hat{r}(t)}{2})^{t}}+\frac{1}{(1+\frac{\hat{r}(T)}{2})^{2T}} &>\\ \frac{C(T)}{2}\sum_{t=1}^{2T}\frac{1}{(1+\frac{\hat{r}(T)}{2})^{t}}+\frac{1}{(1+\frac{\hat{r}(T)}{2})^{2T}}&\end{aligned} \tag{A2.24}$$

Note that the spot rates in the summation on the top line are $\hat{r}(t)$, while those in the summation in the bottom line are all $\hat{r}(T)$.

Now, because the left-hand sides of Equations (A2.22), (A2.23), and (A2.24) are all equal to one, the left-hand side of (A2.23) can replace the left-hand side of (A2.24), that is,

$$\begin{aligned}\frac{C(T)}{2}\sum_{t=1}^{2T}\frac{1}{\left(1+\frac{C(T)}{2}\right)^{t}}+\frac{1}{\left(1+\frac{C(T)}{2}\right)^{2T}} &>\\ \frac{C(T)}{2}\sum_{t=1}^{2T}\frac{1}{\left(1+\frac{\hat{r}(T)}{2}\right)^{t}}+\frac{1}{\left(1+\frac{\hat{r}(T)}{2}\right)^{2T}}&\end{aligned} \tag{A2.25}$$

which implies that $C(T) < \hat{r}(T)$, which was to be proved.

Proposition 6: If $\hat{r}(0.5) > \hat{r}(1.0) > \cdots > \hat{r}(T)$, then $C(T) > \hat{r}(T)$.

Proof: Reverse the inequalities of Equations (A2.24) and (A2.25) in the previous proof to conclude that $C(T) > \hat{r}(T)$, as was to be proved.

APPENDIX TO CHAPTER 3

Returns, Yields, Spreads, and P&L Attribution

A3.1 YIELD TO MATURITY FOR SETTLEMENT DATES OTHER THAN COUPON PAYMENT DATES

Equations (3.6), (3.7), and (3.8) express the relationship between price and yield to maturity when settlement is on a coupon payment date. This appendix generalizes this relationship to other settlement dates. First, because accrued interest is zero when settlement falls on a coupon payment date, the full price in the equations in the text do not include any accrued interest. In this section, however, because settlement can fall on other dates, P is taken to include accrued interest.

Second, the market convention for discounting cash flows that do not occur in regular six-month intervals, using a semiannually compounded rate, is as follows. Let y denote the semiannually compounded yield, and let τ denote the fraction of a semiannual period until the next coupon payment. For example, if the next coupon payment is in one month, taken to mean one-sixth of a semiannual period, then $\tau = 1/6$. By convention, then, the present value of a unit of currency at that time is,

$$\frac{1}{\left(1+\frac{y}{2}\right)^{\tau}} \tag{A3.1}$$

Note that, while reasonably intuitive, this convention cannot really be justified by the logic of compounding conventions. As discussed in the text, $(1 + y/2)^N$ represents the final proceeds of an investment of one unit of currency semiannually compounded N times. There is no such interpretation for an exponent that is not a whole number of semiannual periods. In any case, continuing along these lines, the present value of a unit of currency to be paid after $\tau + i$ semiannual periods is,

$$\frac{1}{\left(1+\frac{y}{2}\right)^{\tau+i}} \tag{A3.2}$$

Finally, then, consider a bond with $2T$ remaining coupon payments of $c/2$, the first of which is paid after τ semiannual periods, the second after $\tau + 1$ semiannual periods, the third after $\tau + 2$ semiannual periods, etc., and the last, along with a principal payment of 100, after $\tau + 2T - 1$ semiannual periods. Its price is given by,

$$P = \frac{\frac{1}{2}c}{\left(1+\frac{y}{2}\right)^{\tau}} + \frac{\frac{1}{2}c}{\left(1+\frac{y}{2}\right)^{\tau+1}} + \cdots + \frac{100+\frac{1}{2}c}{\left(1+\frac{y}{2}\right)^{\tau+2T-1}} \tag{A3.3}$$

$$P = \frac{c}{2}\sum_{t=0}^{2T-1} \frac{1}{\left(1+\frac{y}{2}\right)^{\tau+t}} + \frac{100}{\left(1+\frac{y}{2}\right)^{\tau+2T-1}} \tag{A3.4}$$

$$P = \left(1+\frac{y}{2}\right)^{1-\tau}\left[\frac{c}{y}\left(1 - \frac{1}{\left(1+\frac{y}{2}\right)^{2T}}\right) + \frac{100}{\left(1+\frac{y}{2}\right)^{2T}}\right] \tag{A3.5}$$

where (A3.5) can be derived from Equation (A2.16) in Appendix A2.2.

A3.2 YIELD TO MATURITY AND *EX-POST* RETURNS

For simplicity, this section assumes annual coupons and annual compounding.

Proposition: A T-year coupon bond priced at a yield of y earns y per year over n years if its coupons are all reinvested at y and if the bond's yield at the end of n years is y.

Proof: Let P_0 and P_n be the prices of the bond at time 0 (T years to maturity) and after n years ($T - n$ years to maturity) when its yield is y.

Starting with the definition of yield to maturity,

$$\begin{aligned} P_0 = {} & \frac{c}{(1+y)} + \frac{c}{(1+y)^2} + \cdots + \frac{c}{(1+y)^{n-1}} + \frac{c}{(1+y)^n} \\ & + \frac{c}{(1+y)^{n+1}} + \frac{c}{(1+y)^{n+2}} + \cdots + \frac{100+c}{(1+y)^T} \end{aligned} \tag{A3.6}$$

$$P_0(1+y)^n = c(1+y)^{n-1} + c(1+y)^{n-2} + \cdots + c(1+y) + c$$
$$+ \frac{c}{(1+y)} + \frac{c}{(1+y)^2} + \cdots + \frac{100+c}{(1+y)^{T-n}} \quad \text{(A3.7)}$$

$$P_0(1+y)^n = c(1+y)^{n-1} + \cdots + c(1+y) + c + P_n \quad \text{(A3.8)}$$

$$(1+y)^n - 1 = \frac{c(1+y)^{n-1} + \cdots + c(1+y) + c + P_n - P_0}{P_0} \quad \text{(A3.9)}$$

where (A3.7) simply multiplies both sides of (A3.6) by $(1+y)^n$; (A3.8) recognizes that the second line of (A3.7) is just P_n, the price of the bond after n years, with a remaining maturity of $T-n$ years and a yield of y; and (A3.9) simplifies terms.

By inspection, the right-hand side of Equation (A3.9) is the n-year return of the bond if all coupons are reinvested at y and if the yield after n years is y. Breaking that down, the numerator gives the reinvested coupon payments at a yield of y; plus P_n, the price of the bond after n years if the yield is y; minus P_0 the initial price of the bond at the yield y. Hence, the numerator divided by the initial price equals the n-year return.

But if the right-hand side of the bond is the n-year return under the conditions stated, then, by the left-hand side, that return is equivalent to earning y per year for n years.

Note that the proposition is not an "if and only if," because it is possible – though very unlikely – to have some combination of returns on coupons and a final yield that also result in a holding-period return of y per year.

A3.3 REALIZED FORWARD SCENARIO

For simplicity, this section assumes annual coupons and annual compounding.

Proposition: Under the realized forward scenario, the n-year return of a T-year coupon, with coupon income reinvested at the initial forward rates $f(1), f(2), \ldots f(n)$, is the same as from rolling over a unit of currency at those forward rates.

Proof: Let P_0 and P_n be the prices of the bond at time 0 (T years to maturity) and after n years ($T-n$ years to maturity) under the realized forward scenario.

Starting with the definition of forward rates,

$$
\begin{aligned}
P_0 = {} & \frac{c}{(1+f(1))} + \frac{c}{(1+f(1))(1+f(2))} + \cdots \\
& + \frac{c}{(1+f(1)) \cdots (1+f(n))} \\
& + \frac{c}{(1+f(1)) \cdots (1+f(n+1))} + \cdots \\
& + \frac{100+c}{(1+f(1)) \cdots (1+f(T))}
\end{aligned}
\tag{A3.10}
$$

$$
\begin{aligned}
P_0(1+f(1)) \cdots (1+f(n)) = {} & c(1+f(2)) \cdots (1+f(n)) + \cdots + c \\
& + \frac{c}{(1+f(n+1))} + \cdots \\
& + \frac{100+c}{(1+f(n+1)) \cdots (1+f(T))}
\end{aligned}
\tag{A3.11}
$$

$$
P_0(1+f(1)) \cdots (1+f(n)) = c(1+f(2)) \cdots (1+f(n)) + \cdots + c + P_n \tag{A3.12}
$$

$$
(1+f(1)) \cdots (1+f(n)) - 1 = \frac{c(1+f(2)) \cdots (1+f(n)) + \cdots + c + P_n - P_0}{P_0} \tag{A3.13}
$$

where (A3.11) simply multiplies both sides of (A3.10) by $(1+f(1)) \cdots (1+f(n))$; (A3.12) recognizes that the second and third lines of (A3.11) are P_n, the price of the bond after n years, with discounting under the assumption of realized forwards; and (A3.13) simplifies terms.

By inspection, the right-hand side of Equation (A3.13) is the n-year return of the bond if all coupons are reinvested at the short-term rate of the realized forward scenario and if, after n years, the bond is priced under that scenario. But if the right-hand side is the bond's n-year return under the conditions stated, then, by the left-hand side, that return is equivalent to rolling over a unit of currency at the initial forward rates for n years.

Corollary: Under the realized forward scenario, the one-year return of any coupon bond is the short-term rate, $f(1)$.

Proof: Following the proof of the previous proposition with $n = 1$, Equation (A3.13) becomes,

$$
f(1) = \frac{c + P_1 - P_0}{P_0} \tag{A3.14}
$$

which was to be proved.

APPENDIX TO CHAPTER 4

DV01, Duration, and Convexity

A4.1 DV01, DURATION, AND CONVEXITY OF PORTFOLIOS

Let P^i denote the price of asset i and P the price of a portfolio of those assets. By definition,

$$P = \sum P^i \tag{A4.1}$$

Let y be the single factor generating changes in rates. Then, taking the derivative of both sides of (A4.1) with respect to y,

$$\frac{dP}{dy} = \sum \frac{dP^i}{dy} \tag{A4.2}$$

Then, divide both sides of (A4.2) by −10,000 and apply Equation (4.5) of the text to see that,

$$DV01 = \sum DV01^i \tag{A4.3}$$

Or, in words, the DV01 of a portfolio equals the sum of the individual asset DV01s.

To derive the duration of a portfolio, start from Equation (A4.2), dividing both sides by $-P$,

$$-\frac{1}{P}\frac{dP}{dy} = \sum -\frac{1}{P}\frac{dP^i}{dy} \tag{A4.4}$$

$$-\frac{1}{P}\frac{dP}{dy} = \sum -\frac{P^i}{P}\frac{1}{P^i}\frac{dP^i}{dy} \tag{A4.5}$$

$$D = \sum \frac{P^i}{P} D^i \tag{A4.6}$$

Equation (A4.5) multiplies each term in the summation by one, in the form of P^i/P^i. Equation (A4.6) follows from the definition of duration in the text, Equation (4.11). In words, Equation (A4.6) says that the duration of a portfolio equals the weighted average of the durations of the individual assets, where the weights are the fraction of value of each asset in the portfolio.

The proof for the convexity of a portfolio can be derived along the same lines as the duration of a portfolio. The result, given here without proof, is,

$$C = \sum \frac{P^i}{P} C^i \tag{A4.7}$$

A4.2 ESTIMATING PRICE CHANGE WITH DURATION AND CONVEXITY

Let $P(y)$ be the price of an asset as a function of the single factor that describes changes in rates. Then, a *second-order Taylor approximation* of the price rate function is given by,

$$P(y + \Delta y) \approx P(y) + \frac{dP}{dy}\Delta y + \frac{1}{2}\frac{d^2P}{dy^2}\Delta y^2 \tag{A4.8}$$

Subtracting $P(y)$ from both sides of (A4.8) and denoting the change in price, $P(y + \Delta y) - P$, by ΔP,

$$\Delta P \approx \frac{dP}{dy}\Delta y + \frac{1}{2}\frac{d^2P}{dy^2}\Delta y^2 \tag{A4.9}$$

$$\frac{\Delta P}{P} \approx \frac{1}{P}\frac{dP}{dy}\Delta y + \frac{1}{2}\frac{1}{P}\frac{d^2P}{dy^2}\Delta y^2 \tag{A4.10}$$

$$\frac{\Delta P}{P} \approx -D\Delta y + \frac{1}{2}C\Delta y^2 \tag{A4.11}$$

where (A4.11) – which is Equation (4.16) – follows from the definitions of duration and convexity, that is, from Equation (4.10) and a discrete version of Equation (4.14).

APPENDIX TO CHAPTER 6

Regression Hedging and Principal Component Analysis

A6.1 REGRESSION HEDGES AND P&L VARIANCE

This section proves that i) the regression hedge minimizes the variance of the P&L of the hedged portfolio; and ii) the volatility of the regression-hedged portfolio equals the DV01 of the position being hedged times the standard deviation of the regression residuals.

Begin with least-squares estimation, which finds the parameters $\hat{\alpha}$ and $\hat{\beta}$ to minimize,

$$\sum_t (\Delta y_t - \hat{\alpha} - \hat{\beta}\Delta x_t)^2 \tag{A6.1}$$

To solve this minimization, differentiate (A6.1) with respect to each of the parameters, set each result to zero, and obtain the following two equations,

$$-2\sum_t (\Delta y_t - \hat{\alpha} - \hat{\beta}\Delta x_t) = 0 \tag{A6.2}$$

$$-2\sum_t (\Delta y_t - \hat{\alpha} - \hat{\beta}\Delta x_t)\Delta x_t = 0 \tag{A6.3}$$

These equations can be solved to show that,

$$\hat{\alpha} = \overline{\Delta y} - \hat{\beta}\overline{\Delta x} \tag{A6.4}$$

$$\hat{\beta} = \frac{\sigma_{xy}}{\sigma_x^2} = \frac{\rho\sigma_y}{\sigma_x} \tag{A6.5}$$

where $\overline{\Delta x}$ and $\overline{\Delta y}$ are the sample averages; σ_x and σ_y the standard deviations; σ_{xy} the covariance; and ρ the correlation. The solutions (A6.4) and (A6.5) are not derived step-by-step here, but are easily found by noting that, with N observations, the summary statistics needed are defined as follows,

$$\overline{\Delta x} = \frac{\sum_t \Delta x_t}{N} \tag{A6.6}$$

$$\overline{\Delta y} = \frac{\sum_t \Delta y_t}{N} \tag{A6.7}$$

$$\sigma_x^2 = \frac{\sum_t \Delta x_t^2}{N} - \left(\frac{\sum_t \Delta x_t}{N}\right)^2 \tag{A6.8}$$

$$\sigma_y^2 = \frac{\sum_t \Delta y_t^2}{N} - \left(\frac{\sum_t \Delta y_t}{N}\right)^2 \tag{A6.9}$$

$$\sigma_{xy} = \frac{\sum_t \Delta y_t \Delta x_t}{N} - \frac{\sum_t \Delta y_t}{N}\frac{\sum_t \Delta x_t}{N} \tag{A6.10}$$

$$\rho = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \tag{A6.11}$$

The discussion now turns to minimizing the P&L of the hedged position. That P&L, given in Equation (6.11), is repeated here for convenience,

$$P\&L = -F^{JNJ}\frac{DV01^{JNJ}}{100}\Delta y_t^{JNJ} - F^{30}\frac{DV01^{30}}{100}\Delta y_t^{30} \tag{A6.12}$$

To simplify notation, write the DV01s of the bond positions as,

$$\overline{DV01}^{JNJ} \equiv \frac{F^{JNJ} DV01^{JNJ}}{100} \tag{A6.13}$$

$$\overline{DV01}^{30} \equiv \frac{F^{30} DV01^{30}}{100} \tag{A6.14}$$

Then, with the obvious notations for variance and covariance, the variance of the P&L in (A6.12), denoted $\sigma_{P\&L}^2$, is,[1]

$$\sigma_{P\&L}^2 = \left(\overline{DV01}^{JNJ}\right)^2 \sigma_{JNJ}^2 + \left(\overline{DV01}^{30}\right)^2 \sigma_{30}^2 + 2\overline{DV01}^{JNJ}\,\overline{DV01}^{30}\sigma_{JNJ,30} \tag{A6.15}$$

[1]Recall that for two random variables, w and z, and two constants, a and b, the variance of $aw + bz$ equals $a^2\sigma_w^2 + b^2\sigma_z^2 + 2ab\sigma_{wz}$.

To minimize this variance by choosing the DV01 in the hedging bonds, differentiate (A6.15) with respect to $\overline{DV01}^{30}$, set the result to zero, and solve for $\overline{DV01}^{30}$,

$$2\overline{DV01}^{30}\sigma_{30}^2 + 2\overline{DV01}^{JNJ}\sigma_{JNJ,30} = 0$$

$$\overline{DV01}^{30} = -\overline{DV01}^{JNJ}\frac{\sigma_{JNJ,30}}{\sigma_{30}^2} \tag{A6.16}$$

But, by inspection of Equation (A6.5), the fraction in (A6.16) is just the estimated slope coefficient in a regression of JNJ yields on 30-year Treasury yields. Hence, the regression hedge given in Equations (6.8) or (6.10) minimizes the variance of the P&L of the hedged position.

The minimized P&L variance of the hedged portfolio can be written explicitly by substituting Equation (A6.16) into Equation (A6.15) and rearranging terms,

$$\sigma_{P\&L}^2 = \left(\overline{DV01}^{JNJ}\right)^2\sigma_{JNJ}^2 + \left(\overline{DV01}^{JNJ}\right)^2\frac{\sigma_{JNJ,30}^2}{\sigma_{30}^2}$$

$$-2\left(\overline{DV01}^{JNJ}\right)\left[\overline{DV01}^{JNJ}\frac{\sigma_{JNJ,30}}{\sigma_{30}^2}\right]\sigma_{JNJ,30} \tag{A6.17}$$

$$= \left(\overline{DV01}^{JNJ}\right)^2\sigma_{JNJ}^2\left[1 - \frac{\sigma_{JNJ,30}^2}{\sigma_{30}^2\sigma_{JNJ}^2}\right] \tag{A6.18}$$

$$= \left(\overline{DV01}^{JNJ}\right)^2\sigma_{JNJ}^2[1-\rho^2] \tag{A6.19}$$

where ρ denotes the correlation between changes in the JNJ and Treasury bond yields.

The last step for this section is to show that the variance of the hedged P&L, now given in (A6.19), is equal to the squared DV01 of the bonds being hedged times the variance of the regression residuals. Starting with the definition of the regression residuals in the general regression context of (A6.1), their variance, denoted σ_ϵ^2, can be expressed as follows,

$$\epsilon_t = \Delta y_t - \alpha - \beta\Delta x_t \tag{A6.20}$$

$$\sigma_\epsilon^2 = \sigma_y^2 + \beta^2\sigma_x^2 - 2\beta\sigma_{xy} \tag{A6.21}$$

$$= \sigma_y^2 + \left(\frac{\rho\sigma_y}{\sigma_x}\right)^2\sigma_x^2 - 2\left(\frac{\rho\sigma_y}{\sigma_x}\right)\sigma_{xy} \tag{A6.22}$$

$$= \sigma_y^2(1-\rho^2) \tag{A6.23}$$

But applying Equation (A6.23) to the regression of the JNJ bonds on the 30-year Treasury bonds, and multiplying by the DV01 of the JNJ bonds, gives exactly the right-hand side of Equation (A6.19), which was to be proved.

A6.2 CONSTRUCTION OF PRINCIPAL COMPONENTS

The goal of this section is to illustrate how PCs are constructed with a minimum of mathematics. A slightly more rigorous mathematical treatment is given in Section A6.3. For illustration, this section uses the data from the text on daily, basis-point changes in the five-, 10-, and 30-year swap rates only. The covariance matrix, or the variance-covariance matrix of these rate changes is,

$$\mathrm{V}=\begin{pmatrix}6.46 & 7.71 & 7.10\\ 7.71 & 11.89 & 12.95\\ 7.10 & 12.95 & 16.19\end{pmatrix} \tag{A6.24}$$

The diagonal of the matrix in (A6.24) gives the variances of the three rates, or, taking square roots, the standard deviations. The off-diagonals give the pairwise covariances of rates, from which the correlations can be derived. For example, the volatilities of the five- and 10-year rates are the square roots of 6.46 and 11.89, or 2.54 and 3.45 basis points per day, respectively, and the correlation between them is $7.71/(2.54\times 3.45) = 88.0\%$. Note, in passing, that the sum of the variances is $6.46 + 11.89 + 16.19 = 34.54$, a number that appears again below.

Now consider portfolio weights or *loadings* of −0.5, 1.0, and −0.6 on the five-, 10-, and 30-year rates, respectively. By the properties of variance and covariance, and with the specific covariance matrix (A6.24), the variance of this portfolio, denoted σ_p^2, is,

$$\begin{aligned}\sigma_p^2 &= (-0.5)^2\times 6.46 + 1.0^2\times 11.89 + (-0.6)^2\times 16.19\\ &\quad + 2\times(-0.5)\times(1.0)\times 7.71\\ &\quad + 2\times(-0.5)\times(-0.6)\times 7.10\\ &\quad + 2\times(1.0)\times(-0.6)\times 12.95\\ &= 0.5860^2\end{aligned} \tag{A6.25}$$

Computations like this are more conveniently written with matrix notation. Let the vector of portfolio weights be **w**, which, in the present example, is $\mathbf{w}' = (-0.5, 1.0, -0.6)$, where the apostrophe denotes the transpose. Then, the same variance as computed in Equation (A6.25) can be written as,

$$\mathbf{w'Vw} = \begin{pmatrix}-0.5 & 1.0 & -0.6\end{pmatrix}\begin{pmatrix}6.46 & 7.71 & 7.10\\ 7.71 & 11.89 & 12.95\\ 7.10 & 12.95 & 16.19\end{pmatrix}\begin{pmatrix}-0.5\\ 1.0\\ -0.6\end{pmatrix} \tag{A6.26}$$

Turning to the creation of the PCs, denote the first principal component by the vector of weights $\mathbf{a} = (a_1, a_2, a_3)'$. Then, solve for the elements of **a** by maximizing the variance of this PC, **a'Va**, such that **a'a** = 1. Maximization ensures that, among all the PCs, the first explains the largest fraction of the sum or total variance across all rates. But there has to be some limit on the vector **a**, or the maximization would find portfolios with arbitrarily large variances. Enter the constraint **a'a** = 1, which – along with similar constraints on other PCs – limits the risks of the PCs in a way that equates the sum of the variances of all PCs to the total variance. (See Section A6.3 for more details.) The maximization just described can be solved with the solver in Excel or some other tool to obtain that $\mathbf{a} = (0.3846, 0.6090, 0.6937)'$. The variance of this PC is **a'Va** = 31.4977, which is 91.2% of the total variance of 34.54 given above.

The second principal component, denoted by $\mathbf{b} = (b_1, b_2, b_3)'$, maximizes **b'Vb** such that **b'b** = 1 and **b'a** = 0. This last constraint ensures that the portfolio represented by the second PC is uncorrelated with the portfolio represented by the first. (Again, see Section A6.3 for more details.) Solving this maximization, $\mathbf{b} = (-0.7851, -0.1793, 0.5928)'$. The variance of this PC is **b'Vb** = 2.8626, which is 8.3% of the total variance of 34.54.

Finally, the third PC, denoted by $\mathbf{c} = (c_1, c_2, c_3)'$, satisfies **c'c** = 1, **c'a** = 0, and **c'b** = 0. Solving, $\mathbf{c} = (0.4854, -0.7726, 0.4092)'$. No maximization is needed here because, by construction, this third PC explains all of the remaining total variance. The variance of this PC is **c'Vc** = 0.1797, which is the remaining 0.5% of the total variance of 34.54.

The maximizations just described constrain the sum of squares of the elements of each PC to equal one. But a different scaling turns out to be convenient for interpreting the PCs: multiply each element of a PC by the volatility of that PC. In that case, the sum of squares of the elements of a PC equals its variance. In addition, after this scaling, the elements of each PC can be interpreted as the number of basis points corresponding to a one standard deviation shift in that PC. (Section A6.3 gives a more precise explanation of this point.) In the current example, the volatilities of the three PCs, from their variances computed above, are $\sqrt{31.4977} = 5.6123$, $\sqrt{2.8626} = 1.6919$, and $\sqrt{0.1797} = 0.4239$, respectively. Multiplying the elements of the respective raw PCs by these numbers gives the scaled PCs in

TABLE A6.1 Principal Components of USD LIBOR Swap Rates, from June 1, 2020, to July 16, 2021, Using Only Five-, 10-, and 30-Year Rates. Entries Are in Basis Points.

Term	Level	Slope	Curvature
5-Year	2.158	−1.328	0.206
10-Year	3.418	−0.303	−0.328
30-Year	3.893	1.003	0.173

Table A6.1. It can then be said that a one standard deviation shock of the level PC is a 2.158-basis-point shift in the five-year rate, a 3.418-basis-point shift in the 10-year rate, and a 3.893-basis-point shift in the 30-year rate. The scaled slope and curvature PCs can be interpreted analogously.

A6.3 CONSTRUCTION OF PCs: MATHEMATICAL DETAILS

This section is more precise on a few claims made in the previous section at the cost of some extra mathematics. Let $\mathbf{V}$ denote the 3×3 variance-covariance matrix of rates with elements V_{ij}; let $\mathbf{P}$ denote the 3×3 matrix of principal components, with elements p_{ij}, or, alternatively, with three 3×1 column vectors $\mathbf{p_i}$ corresponding to PC i; let $\mathbf{D}$ denote the 3×3 diagonal matrix with diagonal elements σ_i^2, each equal to the variance of PC i; and let $\mathbf{I}$ denote the 3×3 identity matrix. Then, though not proved here, the construction of the PCs in the previous section guarantees that,

$$\mathbf{V} = \mathbf{PDP'} \tag{A6.27}$$

$$\mathbf{P'P} = \mathbf{PP'} = \mathbf{I} \tag{A6.28}$$

$$\mathbf{P'VP} = \mathbf{P'PDP'P} = \mathbf{D} \tag{A6.29}$$

where (A6.29) follows from (A6.27) and (A6.28).

Lemma 1: The PCs are uncorrelated.

Proof: In terms of its columns, $\mathbf{P} = (\mathbf{p_1}, \mathbf{p_2}, \mathbf{p_3})$. With this, rewrite Equation (A6.29) as,

$$\mathbf{P'VP} = \begin{pmatrix} \mathbf{p_1'Vp_1} & \mathbf{p_1'Vp_2} & \mathbf{p_1'Vp_3} \\ \mathbf{p_2'Vp_1} & \mathbf{p_2'Vp_2} & \mathbf{p_2'Vp_3} \\ \mathbf{p_3'Vp_1} & \mathbf{p_3'Vp_2} & \mathbf{p_3'Vp_3} \end{pmatrix} = \mathbf{D} \tag{A6.30}$$

Because D is diagonal, the numbers $\mathbf{p}_1'\mathbf{V}\mathbf{p}_2$, $\mathbf{p}_1'\mathbf{V}\mathbf{p}_3$, $\mathbf{p}_2'\mathbf{V}\mathbf{p}_3$ are all zero. This means that the pairwise covariances of the PCs are zero, or, equivalently, that the PCs are uncorrelated with each other.

Lemma 2: The variance of rate j equals the sum of the variance of each PC times the square of its jth component. Mathematically,

$$V_{jj} = p_{1j}^2\sigma_1^2 + p_{2j}^2\sigma_2^2 + p_{3j}^2\sigma_3^2 \tag{A6.31}$$

Proof: For $i = 1$, pre-multiply each side of Equation (A6.27) by the 1×3 vector $(1,0,0)$ and post-multiply by the 3×1 vector $(1,0,0)'$. Then,

$$\begin{pmatrix}1 & 0 & 0\end{pmatrix} \mathbf{V} \begin{pmatrix}1\\0\\0\end{pmatrix} = \begin{pmatrix}1 & 0 & 0\end{pmatrix} \mathbf{PDP'} \begin{pmatrix}1\\0\\0\end{pmatrix} \tag{A6.32}$$

Equation (A6.31) then follows by algebra. For $i = 2$, the proof is the same but with the vector $(0,1,0)$ and its transpose, and for $i = 3$ with $(0,0,1)$.

Lemma 3: The sum of the variances of the PCs equals the sum of the variances of the rates.

Proof: Adding together Equations (A6.31) for each j and rearranging terms,

$$\begin{aligned} V_{11} + V_{22} + V_{33} &= p_{11}^2\sigma_1^2 + p_{21}^2\sigma_2^2 + p_{31}^2\sigma_3^2 \\ &\quad + p_{12}^2\sigma_1^2 + p_{22}^2\sigma_2^2 + p_{32}^2\sigma_3^2 \\ &\quad + p_{13}^2\sigma_1^2 + p_{23}^2\sigma_2^2 + p_{33}^2\sigma_3^2 \qquad (A6.33)\\ &= \sigma_1^2(p_{11}^2 + p_{12}^2 + p_{13}^2) \\ &\quad + \sigma_2^2(p_{21}^2 + p_{22}^2 + p_{23}^2) \\ &\quad + \sigma_3^2(p_{31}^2 + p_{32}^2 + p_{33}^2) \qquad (A6.34) \end{aligned}$$

But by Equation (A6.28), the sum of squares of the elements of each PC in the brackets, $\mathbf{p}_i'\mathbf{p}_i$, equals 1, thus proving the lemma.

Lemma 4: Defined a scaled principal component matrix, $\tilde{\mathbf{P}}$, with elements $\tilde{p}_{ij} = \sigma_i p_{ij}$. Then,

$$\sqrt{V_{jj}} = \sqrt{\tilde{p}_{1j}^2 + \tilde{p}_{2j}^2 + \tilde{p}_{3j}^2} \tag{A6.35}$$

Proof: This lemma follows directly from the definition of the $\tilde{p}_{ij}$ and Equation (A6.31).

To understand the significance of Lemma 4, interpret the element $\tilde{p}_{ij}$ as the standard deviation of changes in the *j*th rate, in basis points, due to the scaled *i*th PC. Then, because the PCs are uncorrelated, the standard deviation of the *j*th rate, with contributions from all three scaled PCs, equals the right-hand side of Equation (A6.35). But the left-hand side of the equation is exactly the volatility of the *j*th rate. Hence, taken as a whole, Equation (A6.35) supports the interpretation of the elements of each scaled PCs as one standard deviation shifts in the three rates.

APPENDIX TO CHAPTER 8

Expectations, Risk Premium, Convexity and the Shape of the Term Structure

This appendix proves Equation (8.18) along the lines of Ingersoll (1987).[1] Assume that r, the single, instantaneous rate factor, follows the process,

$$dr = \mu dt + \sigma dw \tag{A8.1}$$

Let P be the full price of some security that depends on r and time. Then, by Ito's lemma,

$$dP = P_r dr + P_t dt + \frac{1}{2} P_{rr} \sigma^2 dt \tag{A8.2}$$

where P_r, P_t, and P_{rr} denote the partial first derivatives with respect to r and t and the second partial derivative with respect to r. Dividing both sides of (A8.2) by P, taking expectations, and defining α_P to be the expected return of the security,

$$\alpha_P dt \equiv E\left[\frac{dP}{P}\right] = \frac{P_r}{P} \mu dt + \frac{P_t}{P} dt + \frac{1}{2} \frac{P_{rr}}{P} \sigma^2 dt \tag{A8.3}$$

Combining (A8.1), (A8.2), and (A8.3),

$$\frac{dP}{P} - \alpha_P dt = \frac{P_r}{P} \sigma dw \tag{A8.4}$$

Because Equation (A8.4) applies to any security, it also applies to some other security Q,

$$\frac{dQ}{Q} - \alpha_Q dt = \frac{Q_r}{Q} \sigma dw \tag{A8.5}$$

[1] Ingersoll, J. (1987), *Theory of Financial Decision Making*, Rowman & Littlefield.

Now consider the strategy of investing one unit of currency in security P and $-P_rQ/PQ_r$ in security Q. From Equations (A8.4) and (A8.5), the return on this portfolio is,

$$\frac{dP}{P} - \frac{P_rQ}{PQ_r}\frac{dQ}{Q} = \alpha_P dt - \frac{P_rQ}{PQ_r}\alpha_Q dt \tag{A8.6}$$

Note that terms with the random variable dw have fallen out of Equation (A8.6). This particular portfolio is chosen, in fact, so as to hedge completely the risk of P with Q. In any case, because the portfolio has no risk, it must earn the instantaneous rate, r_0,

$$\alpha_P dt - \frac{P_rQ}{PQ_r}\alpha_Q dt = \left(1 - \frac{P_rQ}{PQ_r}\right) r_0 dt \tag{A8.7}$$

Rearranging terms,

$$\frac{\alpha_P - r_0}{-P_r/P} = \frac{\alpha_Q - r_0}{-Q_r/Q} \equiv \lambda(r_0, t) \tag{A8.8}$$

Equation (A8.8) says that the expected return of any security above the instantaneous rate divided by its duration with respect to that rate must equal some function λ. This function cannot depend on any characteristic of the security, because (A8.8) is true for all securities. The function may depend, however, on the interest rate factor and time. Rewriting Equation (A8.8), for any security P,

$$E\left[\frac{dP}{P}\right] \equiv \alpha_P dt = r_0 dt + \lambda D dt \tag{A8.9}$$

APPENDIX TO CHAPTER 9

The Vasicek and Gauss+ Models

A9.1 THE VASICEK MODEL IN A BINOMIAL TREE

This section illustrates the implementation of the Vasicek model in a binomial tree. The parameters are $r_0 = 2\%$; $\theta = 11\%$; $k = 0.0165$; and $\sigma = 0.95\%$. The step size is one year. The first stages of construction are shown in Figure A9.1. The starting short-term rate, by definition is 2%. By a discrete time approximation to the dynamics of the risk-neutral process in Equation (9.1) of the text, the expected short-term rate after one year is,

$$r_0 + k(\theta - r_0)dt = 2\% + 0.0165(11\% - 2\%) \times 1 = 2.1485\% \tag{A9.1}$$

Note that, if the time step were one month, instead of one year, the *dt* factor would be 1/12 instead of 1. In any case, adding a volatility of 0.95% up and down around the expectation from (A9.1) gives the date-1 up- and down-states in Figure A9.1.

Because of the mean reversion in the Vasicek model, computing the evolution of the short-term rate over the next date is more complicated than the examples in Chapter 7. More specifically, because the drifts from the two states on date 1 are different from the drift from date 0, the tree does not necessarily recombine. One methodology for constructing a recombining tree is the following, though a full analysis of numerical issues is beyond the scope of this treatment.

Given that the expected short-term rate on date 1 is 2.1485%, as calculated in (A9.1), the expected short-term rate on date 2 is,

$$2.1485\% + 0.0165(11\% - 2.1485\%) \times 1 = 2.2945\% \tag{A9.2}$$

which is the short-term rate assigned to the date-2, state 1 node of Figure A9.1.

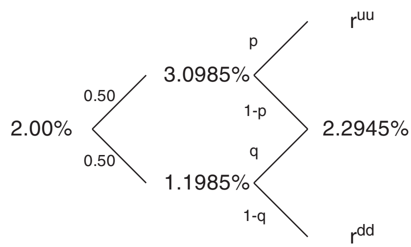

FIGURE A9.1 Binomial Tree Setup for Three Dates of the Vasicek Model.

The next stages of the construction are to find the missing rates and probabilities in Figure A9.1. From the date-1 up-state, the unknowns, r^{uu} and p, must result in the drift and volatility specified by the model. Mathematically, the drift condition is,

$$3.0985\% + 0.0165(11\% - 3.0985\%) = 3.2289\%$$
$$= p \times r^{uu} + (1-p) \times 2.2945\% \quad \text{(A9.3)}$$

and the volatility condition is,

$$0.95\% = \sqrt{p(r^{uu} - 3.2289\%)^2 + (1-p)(3.2289\% - 2.2945\%)^2} \quad \text{(A9.4)}$$

Solving Equations (A9.3) and (A9.4) simultaneously shows that $p = .4917$ and $r_{uu} = 4.1949\%$. These values are given in Figure A9.2 along with solutions to the analogous equations for q and r_{dd}.[1]

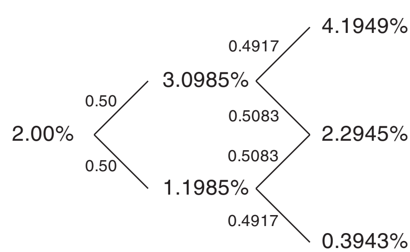

FIGURE A9.2 Binomial Tree Solution for Three Dates of the Vasicek Model.

[1]The values of p and $1 - q$ are the same to four decimal places but are not identically equal.

A9.2 THE GAUSS+ MODEL

A9.2.1 Model Solution

Recall the dynamics of the short rate in cascade form presented in the text, where the cascade form means that each factor mean reverts to another factor, which mean reverts to another factor, in order of persistence. In order to solve the model (where "solving" means solving for the mapping from factors to forward rates), it will be convenient to work with the factors written in reduced form (where "reduced form" means that each factor mean reverts to a constant). After we have solved the model, we will write it back in cascade form in order to proceed to the estimation of its parameters. It will be convenient to partition the parameter vector $P = (\alpha_s, \alpha_m, \alpha_l, \sigma_l, \sigma_m, \rho, \mu)$ in three blocks as $\alpha = (\alpha_s, \alpha_m, \alpha_l)$ and $\sigma = (\sigma_l, \sigma_m, \rho)$ and μ. In the reduced form, we have $r_t = \mu + \mathbf{1}'X_t$ and the relationship between the reduced form expression and cascade form expression of the factors is,

$$x_t = A(\alpha)X_t + \mu \tag{A9.5}$$

where,

$$A(\alpha) = \begin{bmatrix} 1 & 1 & 1 \\ 0 & \frac{\alpha_r - \alpha_m}{\alpha_r} & \frac{\alpha_r - \alpha_l}{\alpha_r} \\ 0 & 0 & \frac{(\alpha_r - \alpha_l)(\alpha_m - \alpha_l)}{\alpha_r} \end{bmatrix} \tag{A9.6}$$

The dynamics of the reduced form factors are then given by,

$$dX_t = -diag(\alpha)X_t dt + A(\alpha)^{-1}\Omega dW_t \tag{A9.7}$$

where for convenience we have appended a zero to the two-dimensional random process dW_t and,

$$\Omega(\sigma) = \begin{bmatrix} 0 & 0 & 0 \\ \rho\sigma_m & \sqrt{1-\rho^2}\sigma_m & 0 \\ \sigma_l & 0 & 0 \end{bmatrix} \tag{A9.8}$$

Let $P_t(\tau)$ denote the zero coupon bond price at time t with time-to-expiry τ. Taking the usual expectation of exponential of future short rate path conditional on the factors expressed in reduced form, one can show that,

$$P_t(\tau) = E^{Q}\left(e^{-\int_0^\tau r_s\, ds}\right) = exp(-y_t(\tau)\tau) \tag{A9.9}$$

where the yield of a zero coupon bond with maturity τ at time t is given by,

$$y_t(\tau) = \mu - C(\tau, \alpha, \sigma) + B(\tau, \alpha)X_t \tag{A9.10}$$

where $B(\tau, \alpha)$ is a three-dimensional vector with $B_i(\tau) = (1 - exp(-\alpha_i\tau))/\alpha_i\tau$ for $i = 1,2,3$, and the term $C(\tau, \alpha, \sigma)$ can be written as,

$$C(\tau, \alpha, \sigma) = \sum_{i=1}^{3}\sum_{j=1}^{3} \frac{\sigma_{ij}}{2\alpha_i\alpha_j}\left(1 - B_i(\tau) - B_j(\tau) - \frac{1 - exp(-(\alpha_i + \alpha_j)\tau)}{(\alpha_i + \alpha_j)\tau}\right) \tag{A9.11}$$

where σ_{ij} stands for the *ith* row and *jth* column of matrix, $A(\alpha)^{-1}\Omega(\sigma)\Omega(\sigma)'$ $A(\alpha)^{-1}$ with,

$$A(\alpha)^{-1} = \begin{bmatrix} 1 & -\frac{\alpha_r}{\alpha_r - \alpha_m} & \frac{\alpha_r\alpha_m}{(\alpha_r-\alpha_m)(\alpha_r-\alpha_l)} \\ 0 & \frac{\alpha_r}{\alpha_r-\alpha_m} & \frac{\alpha_r\alpha_m}{(\alpha_r-\alpha_m)(\alpha_m-\alpha_l)} \\ 0 & 0 & \frac{\alpha_r\alpha_m}{(\alpha_r-\alpha_l)(\alpha_m-\alpha_l)} \end{bmatrix} \tag{A9.12}$$

Mapping reduced form factors in (A9.10) to factors in cascade form, we finally have,

$$y_t(\tau) = \mu(1 - \Upsilon(\tau, \alpha)\mathbf{1}) - C(\tau, \alpha, \sigma) + \Upsilon(\tau, \alpha)x_t \tag{A9.13}$$

where **1** is a three-dimensional column vector of ones and we have defined,

$$\Upsilon(\tau, \alpha) = B(\tau, \alpha)A(\alpha)^{-1} \tag{A9.14}$$

where $\Upsilon(\tau, \alpha) = (\Upsilon_s(\tau, \alpha), \Upsilon_m(\tau, \alpha), \Upsilon_l(\tau, \alpha),)$ where Υ_s, Υ_m and Υ_l represent the partial derivatives (a.k.a. loadings) of the changes of the zero coupon yield of maturity τ on the short, medium and long rate factors. A similar representation for the (continuously compounded) forward rate of tenor τ' as an affine function of the factors in cascade form can be obtained as follows,

$$f_t(\tau) = \mu(1 - \Upsilon'(\tau, \alpha, \tau')\mathbf{1}) + \Upsilon'(\tau, \alpha, \tau')x_t - C'(\tau, \alpha, \sigma, \tau') \tag{A9.15}$$

where,

$$\Upsilon'(\tau, \alpha, \tau') = (B(\tau + \tau', \alpha) - B(\tau, \alpha))A(\alpha)^{-1}/\tau' \tag{A9.16}$$

$$C'(\tau, \alpha, \sigma, \tau') = (C(\tau + \tau', \alpha, \sigma) - C(\tau, \alpha, \sigma, \tau'))/\tau' \tag{A9.17}$$

Expression (A9.15) can be interpreted as follows: The time t forward rate with maturity τ and tenor τ' can be decomposed into a component that represents the risk-neutral expectation of the τ' maturity yield prevailing at time $t + \tau$ (first two terms of the right-hand side) minus a component that represents the convexity advantage of receiving (i.e., being long) rates (third term of the right-hand side). The first on the right-hand side component can be shown to be null for $\tau = 0$ and converges to μ as τ grows toward infinity. For maturities between zero and infinity, the risk-neutral expectation of the zero yield prevailing at that maturity will depend on the value that the factors take at time t, namely, $x_t = (r_t, m_t, l_t)$, and this is what the second term on the right-hand side captures. We can also think of the sum of the first two components on the right-hand side as the fair futures rate at time t for a mark-to-market futures contract on the zero coupon yield that will prevail at time $t + \tau$. The shape of the slopes $\Upsilon'(\tau, \alpha, \tau')$, evaluated at estimated parameter values, are displayed in Figure 9.7.

The last term on the right-hand side of (A9.15) is null for $\tau = 0$ and grows at speed equal to the square of maturity τ. This convexity term drives a wedge between the futures rate and the forward rate. Imagine that an investor can bet on the yield that will prevail at time $t + \tau$ in two different contracts: the first is a daily-settled futures contract. Any increase (decrease) in the market expected future rate will be paid (received) today by the long (short) investor. The second is a standard forward contract. Any increase (decrease) in the market expected future rate will be paid (received) by the long (short) investor not today, but at time $t + \tau$. Now, if the futures rate and the forward rate were equal, the long investor would be at an advantage in transacting via the forward contract rather than the futures contract: when the market future rate expectation increases (decreases) their future losses (gains) will be discounted to today at a relatively higher (lower) rate, hence yielding a relatively smaller (larger) present value loss (gain). Therefore, to prevent an arbitrage between forward and futures rates, today's forward rate should be somewhat lower than today's futures rate. The difference between the two is what we call the *convexity correction* or *convexity advantage term*, that is, the last term on the right hand side.

A9.2.2 A Practical Estimation Method

There exists a large literature on estimating gaussian term structure models, of which the Gauss+ model is a particular case. The methods typically rely on maximum-likelihood estimation procedures which are not easy to implement in general. In the next section we discuss a procedure that is fast and intuitive for the purpose of finding reasonable values for the Gauss+ parameter vector $P = (\alpha, \sigma, \mu)$. Given this parameter vector, and for any time t, we will then extract the factors $x_t = (r_t, m_t, l_t)$ in order to have the model zero yields be close to the market zero yields. The short rate r_t will not need to be filtered because it will be set equal to the observed short policy rate.

In order to estimate parameters α, σ, μ, we first obtain a time series of zero coupon bond prices (discount factors) at different maturities. These can be obtained by using an external curve construction bootstrap method using bonds or par swap rates. For the purpose of fitting the Gauss+ model, it does not matter whether these discount factors come from a swap curve (index or discount curves) or from a bond curve. Here we will use the time series of zero coupon bonds published by the New York Fed, which is based on applying the well-known Nelson-Siegel-Svensson smooth curve construction procedure to US Treasury notes and bonds.[2]

There is always a trade-off when it comes to selecting the sample for the estimation of the model. Longer samples will lead to more robust estimates in a statistical sense but will also reflect different market conditions from the ones in which we intend to use the model. We have found that, in practice, a good compromise is to use a sample size of eight years and a decay factor of 0.8. (This means that in the loss-functions associated with the optimization problems that follow we will give a weight of 1 to the last observation, a weight of 0.8 to an observation that is one year old relative to the last observation, a weight of 0.64 to an observation that is two years old relative to the last observation, and so forth.) Our sample thus begins on January 5, 2014, and ends on January 21, 2022.

While the Gauss+ model involves three factors, only two are genuine factors as we will equal the short rate to a given observed short rate, which tends to change at predictable dates. For this reason, we will take the short rate r_t to be the fed funds target. While taking the general collateral repo rate would be closer to the actual overnight funding rate for US Treasury bonds, this series also exhibits spikes that reflect circumstantial funding conditions and are generally unrelated to monetary policy, and hence unrelated to expectations of future paths of the short rate.

In what follows we will denote by Y the $T \times N$ matrix of observed zero coupon yields for T periods and N maturities, and by Y_t the vector of N maturity yields observed at time t. Similarly, we will denote by y the vector of model implied zero coupon yields for T periods and N maturities. y is a function of x, the $T \times 3$ vector of factors in cascade form. In addition, let and y_t, as described in (A9.13), stand for the vector of yields for the N maturities as a function of the cascade form factors at time t, namely, $x_t = (r_t, m_t, l_t)$. Finally, let $y_t(\tau)$ stand for the yield of maturity τ at time t, as a function of the factors and the parameter vector P.

Estimation involves one step for data preparation, and three steps for parameter optimization. Data preparation consists of *netting* the effect of the (observed) short rate factor from the observed zero coupon yields. This works as follows. Take a candidate parameter value for α_r. Then, given the structure of $\Upsilon(\tau, \alpha)$, subtract $\Upsilon_s(\tau, \alpha) \times r_t$ from both sides of (A9.13). With an abuse of notation, we will denominate $y_t(\tau)$ the zero coupon yield netted

[2]See Gürkaynak, R., Sack, B., and Wright, J. (2006), "The US Treasury Yield Curve: 1961 to the Present," and https://www.federalreserve.gov/data/nominal-yield-curve.html

out of the short rate, and we will drop the short rate r_t from the vector of factors in cascade form, so from now on we have $x_t = (m_t, l_t)$. Incidentally, one could show that $\Upsilon_s(\tau, \alpha)$ depends only on parameter α_r.

The next step consists of inverting (A9.13) and expressing the factors x_t as a linear function of the observed two- and 10-year yields (henceforth 2y and 10y). We will assume that the yields of these two maturities are priced exactly by the model, unlike other maturities. Denote these benchmark yields at time t as $yb_t = (y_t(2), y_t(10))$. We can then invert (A9.13) for the 2y and 10y maturity and express the factors in cascade form as linear functions of the vector of benchmark yields. Then we replace the resulting expression for the factors in (A9.13). Finally, write the resulting expression of benchmark yields as a function of factors, in changes,

$$\Delta yb_t(\tau) = \Upsilon_b(\alpha)\Delta x_t \tag{A9.18}$$

where $\Upsilon_b(\alpha)$ stands for a matrix formed by the vectors (A9.14) corresponding to maturities $\tau = 2{,}10$, and dropping the column. Now, solving for Δx_t and plugging the resulting expression into equation (A9.13) written in changes, we obtain a linear expression relating the yield changes at any maturity to the yield changes of the two benchmark maturities, where the slopes are a nonlinear function of just the parameter α,

$$\Delta y(\tau) = \Upsilon(\alpha)\Upsilon_b(\alpha)^{-1}\Delta yb_t(\tau) \tag{A9.19}$$

we can then compute an estimate for α by solving,

$$min_\alpha \,\|\,\Upsilon(\alpha)\Upsilon_b(\alpha)^{-1} - \hat{\beta}\,\| \tag{A9.20}$$

where $\|\|$ stands for the L2 norm and $\hat{\beta}$ is the ordinary least square estimate of regressing $\Delta y(\tau)$ onto $\Delta yb_t(\tau)$, namely,

$$\hat{\beta} = (\Delta yb'\Delta yb)^{-1}\Delta yb'\Delta y \tag{A9.21}$$

Figure 9.5 in the text shows the estimated OLS parameters of regressing yield changes at different maturities on the changes in the two- and 10-year yield. It also shows the values of the model slopes $\Upsilon(\alpha)\Upsilon_b(\alpha)^{-1}$, evaluated at the optimal solution of (A9.20). The parameter α allows one to examine the impact of a change in the factors r_t, m_t, l_t on the instantaneous forward rates and yields. This allows a clear interpretation of each of the factors: the medium rate m_t has maximum impact around the 2y to 3y maturities; hence it can be interpreted as a monetary policy factor (as has been argued in the text). On the other hand, the long factor exhibits maximum impact around the 7y forward maturity (and 15y for zero yield maturity).

Armed with the estimated parameter $\hat{\alpha}$ that solves (A9.20), we proceed to estimate the vector σ that minimizes the distance between model implied yield volatilities and realized volatilities of constant maturity yields by solving,

$$min_{\sigma} \parallel \Upsilon_b(\hat{\alpha})\Omega(\sigma)\Omega(\sigma)'\Upsilon_b(\hat{\alpha})' - diag(\Delta y'\Delta y)252/T \parallel \quad (A9.22)$$

where Δy is a $T \times N$ vector of yield changes for all maturities. Figure 9.6 in the text shows the fitted zero yield volatilities, versus the volatilities computed directly from observed, constant maturity yield changes.

Finally, using $\hat{\alpha}$ that solves (A9.20) and $\hat{\sigma}$ that solves (A9.22), we determine the parameter μ by minimizing the sum of squares of yield fitting errors, namely,

$$min_{\mu} \sum_{t=1}^{T} \|Y_t - y_t\| \quad (A9.23)$$

where y_t is the vector of model yields for all maturities, at time t. The estimated parameters are shown in Table 9.1 in the text.

Once the parameter vector P has been estimated, we can solve for the factors (m_t, l_t) to ensure that the model fits exactly the two- and 10-year forward rates (with tenor one year) on each date. (Filtering methods such as least-squares or Kalman filtering could be employed to extract fitted factors, but we find exact fitting of two points in the curve is preferable in practical trading applications.) We show the fitted factors in Figure 9.8 in the text, along with the 2y forward rate.[3]

Note that in the text we plot the affine function of l_t, $L(l_t)$ below rather than l_t itself, so that we can interpret the derived long factor as an approximation of the expectation of what the short rate will be in 10 years' time. This is necessary because the extreme persistence (or large half-life, or low mean-reversion parameter) of the factor l_t makes the interpretation of the fitted level of l_t less intuitive. As you can see, the long factor closely tracks the 10y forward rate from before (the difference between the two is explained by the convexity correction adjustment),

$$L(l_t) = \mu(1 - e^{-10\alpha_l}) + l_t e^{-10\alpha_l} \quad (A9.24)$$

The time series properties of the model can be described by graphing its fitted factors over time. As mentioned already, the short-term rate is set each day to the fed funds target rate, and the medium- and long-term factors

[3]While the sample we used for parameter estimation starts in January 2014, we expanded the period for which we extract the filtered factors backward to create Figure 9.8 – holding the estimated parameters constant – in order to facilitate interpretation.

are set so as to match the model and market two and 10 years forward. Figure 9.8 in the text graphs these empirically recovered market factors from January 2007 to January 2022.[4]

A9.2.3 Implied Risk Premia Calculation

Denote by $P(t, \tau)$ the price at time t of a zero coupon bond that matures at time τ and let $p(t, \tau) = log(P(t, \tau))$. Note that here τ stands for a fixed date in the future, and not a date interval.

We will assume henceforth that only the risk of the long factor is priced. This is reasonable for the application we have in mind, which will involve extracting expectations about the 10-year maturity and, at that maturity, the loadings of the forward relative to the short and medium rate are negligible. Now, applying Ito's lemma to the zero coupon bond price as an exponential affine function of the cascade form factors (A9.9) using (A9.13), and passing to the true measure, we can write the dynamics of the instantaneous return of a zero coupon bond as follows,

$$\frac{dP(t,\tau)}{P(t,\tau)} = (r_t + \lambda_t(\tau - t)\Upsilon_3(\tau - t, \alpha)\sigma_l)dt - (\tau - t)\Upsilon(\tau - t, \alpha)\Omega dW_t^* \quad \text{(A9.25)}$$

where the loadings vector $\Upsilon(\tau, \alpha)$ was defined in (A9.14), and $\Upsilon_3(\tau, \alpha)$ stands for its last element, i.e., the loading of a zero coupon yield with maturity $\tau - t$ on the long rate factor. Also, Ω was defined in (A9.8) and W_t^* stands for the Wiener process under the true measure.

The term multiplying dt on the right-hand side of (A9.25) has a clear interpretation: the expected return at time t of holding τ maturity zero coupon bond from time t to time $t + dt$ is equal to the riskless rate prevailing at time t, plus a risk-premium term that is equal to the duration of the zero coupon bond (its maturity) times the loading of the yield of a zero coupon bond with maturity $\tau - t$, times the volatility of the long rate, times the price of risk prevailing at time t, namely, λ_t. We will assume that the price of risk does change over time. We will not specify its dynamics, but we will assume that λ_t is a very persistent process, so that $E_t(\lambda_{t+\Delta t}) \approx \lambda_t$ for $\Delta t = 1$, i.e. one price of risk next-year is expected to be equal to this year's. If we changed measure to the risk-neutral measure, this term would disappear from the right-hand side of (A9.25), that is, the expected instantaneous return of our zero coupon bond would equal the riskless rate.

---

[4]The model is estimated using data from January 2014, but the resulting parameters are used to extract model factors back through 2007. Also, instead of the long-term factor itself, the figure graphs the long-term factor shifted forward 10 years. This allows the series to be more easily interpreted as an approximation for expectations of the short-term rate in 10 years.

Recall as well the definition of the forward rate at time t with maturity τ and tenor $\Delta\tau$: $f_t(\tau) = (p(t,\tau) - p(t,\tau+\Delta\tau))/\Delta\tau$ – we will omit the tenor as an explicitly argument of $f_t()$. Now, consider the following strategy: At time t, buy one zero coupon bond with maturity $\tau + \Delta\tau$, and simultaneously sell one zero coupon bond with maturity τ. First, we claim that the return of this strategy is equal to the forward rate with tenor $\Delta\tau$ prevailing at time t, minus the $\Delta\tau$ maturity yield prevailing at time τ. To see this, denote by R_t^τ the cumulative realized return between time t and time τ. We then have,

$$\begin{aligned} R_t^\tau &= (p(\tau,\tau+\Delta\tau) - p(t,\tau+\Delta\tau)) - (p(\tau,\tau) - p(t,\tau)) \\ &= (p(t,\tau) - p(t,\tau+\Delta\tau)) - (0 - p(t,\tau)) \\ &= (f_t(\tau) - f_\tau(\tau))\Delta\tau \end{aligned} \tag{A9.26}$$

where $f_\tau(\tau)$ is equal to the zero coupon yield with maturity $\Delta\tau$ prevailing at time τ which, for a sufficiently small $\Delta\tau$, approximately equals the spot rate prevailing at time τ. Hence, the expected return (under the true measure) of this strategy at time t is equal to the forward rate at time t, minus the expected short rate at time τ. Let's keep this result aside for a moment.

We will now compute the expected return of this strategy (again, under the true measure). Note that the expected return of the long and short side cancel out for entire the holding period except for the segment $(t, t+\Delta\tau)$ and $(\tau - \Delta\tau, \tau)$, because in the model the risk premia depends on time to maturity only. Applying Ito's lemma to (A9.25), canceling the short rate, integrating and taking expectations, and dismissing the contribution of the segments $(\tau - \Delta\tau, \tau)$ because only the long factor is priced, we get,

$$\begin{aligned} E_t(R_t^\tau) = \int_0^{\Delta\tau} & \lambda_t(\tau+\Delta\tau-t-s)\Upsilon_3(\tau+\Delta\tau-t-s,\alpha)\sigma_l \\ & - \frac{1}{2}(\tau+\Delta\tau-t-s)^2\Upsilon(\tau+\Delta\tau-t-s,\alpha)\Omega\Omega'\Upsilon(\tau+\Delta\tau-t-s,\alpha)'ds \\ = \lambda_t & RP(t,\tau,\Delta\tau) \end{aligned} \tag{A9.27}$$

Note that the expected risk premium (A9.27) is the product of the yet unknown λ_t and an *amount of risk* term $RP(t,\tau,\Delta\tau)$ that can be easily computed. As we mention in the text, in order to imply λ_t from the observed forward rates and the parameters of the model, we will assume that there is a maturity τ such that the (true) expectation of what the short rate will be at maturity τ is equal to the (true) expectation of the short rate will be at any maturity $\tau' > \tau$. With this assumption, we can specialize this reasoning to two long maturities $\tau' > \tau$, subtract the resulting equations and solve for the price of risk λ_t as follows,

$$\lambda_t = \frac{f_t(\tau') - f_t(\tau)}{RP(t,\tau',\Delta\tau)/\Delta\tau - RP(t,\tau,\Delta\tau)/\Delta\tau} \tag{A9.28}$$

Armed with an estimate for λ_t, we can then take expectations on both sides of the equation (A9.26) at maturity $\tau - \Delta\tau$ and use (A9.27) to solve for the expected rate for a long enough maturity $\tau - \Delta\tau$. We finally get,

$$E_t(r_\tau) = f_t(\tau) - \lambda_t RP(t, \tau, \Delta\tau) \tag{A9.29}$$

To get an intuition for (A9.29), say that the price of risk λ_t that we solved before is 0.09. The loading of the 10-year zero coupon bond yield on the long factor is 0.7, and the volatility of the long factor is about 100 basis points, then the 10-year risk premium would be approximately $0.09 \times 0.01 \times 10 \times 0.7 = 63$bps. The convexity advantage term for the 10-year rate is about 24bps. These numbers are very close to the ones obtained by evaluating the model at estimated parameters, and 0.09 is the average price of risk in the sample.[5] If we assume that the 10-year forward rate is 3%, the implied expected one-year rate, nine years rate under the true measure would then be $0.03 + 0.0024 - 0.0063 = 2.61\%$

We computed the price of risk in the manner described already for using the 14- and 15-year forwards to extract λ for every day in our sample, and then using this estimate to imply the expectation of the 10-year rate using the 10-year forward rate. Results are not sensitive to using other, longer maturities.

The implied 10-year rate expectation computed in this manner is plotted in the text against an estimate of the long run expected short rate obtained outside the model, just by just adding the real rate forecast produced by the Cleveland Fed, to a measure of long run inflation, which we took as the average of the 10-year inflation rate forecast produced by the Cleveland Fed, the 10-year forecast produced by the ATSIX model of the Philadelphia Fed, and an estimate of trend inflation obtained with a simple exponential moving average rule with a decay parameter of 0.987, a value used elsewhere in the literature.[6] Other procedures to assess long rate expectations are possible (for example, the Survey of Market Participants conducted by the New York Fed several times a year, or the long-standing Bluechip Financial Forecast survey).

[5]This value may look low for the usual estimates of Sharpe ratio obtained for risky assets. Note, however, that bonds prices have been negatively correlated with risky assets for the last decades, and thus a lower price of interest rate risk relative to risky asset in positive supply (e.g., the stock market), even negative at times, should be expected.

[6]See for example, Cieslak, A., and Povala, P. (2015), "Expected Returns in Treasury Bonds," *Review of Financial Studies* 28(10).

APPENDIX TO CHAPTER 11

Note and Bond Futures

A11.1 FORWARD DROP APPROXIMATELY EQUALS CASH CARRY

Define the following notation, all for 100 face amount of a bond:

- p_0: flat price for spot settlement
- c: annual coupon payment
- d_1: number of days from spot settlement to the coupon payment date
- d_2: number of days from the coupon payment date to forward settlement
- $d = d_1 + d_2$: number of days from spot to forward settlement
- AI_0, AI_d: accrued interest as of spot and forward settlements, respectively
- r: repo rate from spot to forward settlement
- $p_0(d)$: flat price for forward settlement

With this notation, the forward price of the bond, based on the discussion in the text, can be written as,

$$p_0(d) = \left[(p_0 + AI_0)\left(1 + \frac{rd_1}{360}\right) - \frac{c}{2}\right]\left(1 + \frac{rd_2}{360}\right) - AI_d \tag{A11.1}$$

$$p_0 - p_0(d) \approx \frac{c}{2} + AI_d - AI_0 - (p_0 + AI_0)\frac{rd}{360} \tag{A11.2}$$

Equation (A11.2) follows from Equation (A11.1) by dropping terms that are small, because they represent interest-on-interest. In particular, the following approximations are made: $(1 + rd_1/360)(1 + rd_2/360) \approx (1 + rd/360)$; and $(c/2)(1 + rd_2/360) \approx c/2$.

A11.2 FORWARD VERSUS FUTURES PRICES IN A TERM STRUCTURE MODEL

This section highlights the differences between pricing forward and futures contracts in a term structure model. To focus on this difference, it is assumed here that the futures contract has one delivery date and one deliverable bond and that its conversion factor is one. Section A11.4 discusses the pricing of delivery options in a term structure model. Also, for simplicity, it is assumed here that there are no intermediate coupon dates between the spot and forward settlement dates.

Using the methodology of Chapter 7, begin with a recombining, binomial tree with three dates, labeled 0, 1, and 2, where the forward and futures contracts expire on date 2. Set the following notation:

- r_0, r_1^u, and r_1^d: the initial one-period rate, and the one-period rates on date 1, in the up and down states, respectively.
- P_0, P_2^{uu}, P_2^{ud}, and P_2^{dd}: the initial full price of the underlying bond, and the date-2 full prices in the "up–up," "up–down," and "down–down" states, respectively.
- p_0, p_2^{uu}, p_2^{ud}, and p_2^{dd}: flat prices at the indicated dates and states.
- $P_0(2)$: the initial full forward price.
- $p_0(2)$: the initial flat forward price.
- F_0, F_1^u, F_1^d: the initial futures price, and the date-1 futures prices in the up and down states, respectively.
- q, $1 - q$: the risk-neutral transition probabilities on all dates to move up and down from the current state, respectively.

By the definition of a risk-neutral tree, the initial price of the bond can be computed as,

$$P_0 = \frac{1}{1+r_0}\left[q \times \frac{qP_2^{uu} + (1-q)P_2^{ud}}{1+r_1^u} + (1-q) \times \frac{qP_2^{ud} + (1-q)P_2^{dd}}{1+r_1^d}\right] \tag{A11.3}$$

or, rearranging terms,

$$\begin{aligned} P_0 = {} & q^2 \frac{P_2^{uu}}{(1+r_0)(1+r_1^u)} + q(1-q)\frac{P_2^{ud}}{(1+r_0)(1+r_1^u)} \\ & + q(1-q)\frac{P_2^{ud}}{(1+r_0)(1+r_1^d)} + (1-q)^2 \frac{P_2^{dd}}{(1+r_0)(1+r_1^d)} \end{aligned} \tag{A11.4}$$

Each term of Equation (A11.4) is the probability of a particular path through the tree times the discounted value of the bond price at the end of that path, where discounting is done using the short-term rates along the path. Taking all terms together, therefore, P_0 is the expected discounted bond price. More generally, then, letting n be the number of periods, P_n the random full price of the bond at period n, and r_i the random short-term rate over period i to $i + 1$,

$$P_0 = E\left[\frac{P_n}{\prod_{i=0}^{n-1}(1+r_i)}\right] \tag{A11.5}$$

The discount factor to period n, $d(n)$, is just a special case of (A11.5) when the terminal price equals one in all states. Hence,

$$d(n) = E\left[\frac{1}{\prod_{i=0}^{n-1}(1+r_i)}\right] \tag{A11.6}$$

Following the discussion in the text, the forward full price of the bond is just the future value, to the forward delivery date, of the full spot price of the bond. Therefore, using Equations (A11.5) and (A11.6),

$$P_0^{fwd} = \frac{P_0}{d(n)} \tag{A11.7}$$

Turning to the futures price, start again with the recombining, risk-neutral binomial tree. At expiration on date 2, the futures price equals the flat bond price. (At contract expiration, the bond is delivered at the futures price plus accrued interest.) In the up state of date 1, the futures price is F_1^u, which means that the daily settlement payment for a long position of one contract on date 2 is $p_2^{uu} - F_1^u$ in the up-up state and $p_2^{ud} - F_1^u$ in the up-down state. But because the value of a futures contract initiated at any date is zero, the expected discounted value of the daily settlement payments from the date 1 up state must be zero. Hence,

$$\frac{q(p_2^{uu} - F_1^u) + (1-q)(p_2^{ud} - F_1^u)}{1 + r_1^u} = 0 \tag{A11.8}$$

And, solving for the futures price,

$$F_1^u = qp_2^{uu} + (1-q)p_2^{ud} \tag{A11.9}$$

Analogously, the futures price in the date 1 down state is,

$$F_1^d = qp_2^{ud} + (1-q)p_2^{dd} \tag{A11.10}$$

On date 0, the futures price is F_0, and the expected discounted value of the date 1 settlement payment must be zero. Hence,

$$\frac{q(F_1^u - F_0) + (1-q)(F_1^d - F_0)}{1+r_0} = 0 \tag{A11.11}$$

Or,

$$F_0 = qF_1^u + (1-q)F_1^d \tag{A11.12}$$

Finally, substituting Equations (A11.9) and (A11.10) into (A11.12),

$$F_0 = q^2 p_2^{uu} + 2q(1-q)p_2^{ud} + (1-q)^2 p_2^{dd} \tag{A11.13}$$

Hence, in a risk-neutral tree, the futures price is the expected value of the bond price on the contract expiration date. More generally,

$$F_0 = E[p_n] \tag{A11.14}$$

This result is proved more formally in Section A16.5.

A11.3 THE FUTURES-FORWARD DIFFERENCE

Recall that for any two random variables, X and Y, their covariance can be written as,

$$Cov(X,Y) = E[XY] - E[X]E[Y] \tag{A11.15}$$

Applying this to two random variables from the previous section, namely, P_n and $\prod_{i=0}^{n-1}(1+r_i)$,

$$Cov\left(P_n, \frac{1}{\prod_{i=0}^{n-1}(1+r_i)}\right) = E\left[\frac{P_n}{\prod_{i=0}^{n-1}(1+r_i)}\right] - E[P_n]E\left[\frac{1}{\prod_{i=0}^{n-1}(1+r_i)}\right] \tag{A11.16}$$

Let AI_n be the accrued interest on the bond as of date n. Because accrued interest is not a random variable, $E[P_n] = E[p_n] + AI_n$. Next, substitute the definitions of P_0, $d(n)$, and F_0 from Equations (A11.5), (A11.6), and (A11.14) into Equation (A11.16) to see that,

$$Cov\left(P_n, \frac{1}{\prod_{i=0}^{n-1}(1+r_i)}\right) = P_0 - (F_0 + AI_n)d(n) \tag{A11.17}$$

Rearranging terms, substituting in the forward price from Equation (A11.7), and recognizing that $P_0^{fwd} = p_0^{fwd} + AI_n$,

$$
\begin{aligned}
F_0 + AI_n &= \frac{P_0}{d(n)} - \frac{1}{d(n)} Cov\left(P_n, \frac{1}{\prod_{i=0}^{n-1}(1+r_i)}\right) \\
&= P_0^{fwd} - \frac{1}{d(n)} Cov\left(P_n, \frac{1}{\prod_{i=0}^{n-1}(1+r_i)}\right) \\
F_0 &= p_0^{fwd} - \frac{1}{d(n)} Cov\left(P_n, \frac{1}{\prod_{i=0}^{n-1}(1+r_i)}\right) \qquad \text{(A11.18)}
\end{aligned}
$$

Finally, because bond prices are negatively correlated with interest rates, the covariance term in Equation (A11.18) is negative and, therefore,

$$F_0 < p_0^{fwd} \qquad \text{(A11.19)}$$

The futures price is less than the forward price.

A11.4 FUTURES DELIVERY OPTIONS IN A TERM STRUCTURE MODEL

Having set up a term structure model in the form of a tree, the quality option can be priced as follows. Start at the delivery date. At each node, compute the ratio of the price to conversion factor for each deliverable bond. Find the bond with the minimum ratio and set the futures price equal to that ratio. Then, given these terminal values of the futures price, prices on earlier dates can be computed along the lines described in Section A11.2. In the case of the timing, end-of-month, and wild-card options, which are American-style options exercisable by the short, the futures price at each node in the tree is the minimum of the futures price with the option exercised and with the option not exercised.

The previous paragraph assumes that the prices of the bonds are available on the tree as of the last delivery date. These bond prices can be computed in one of two ways. If a model with a closed-form solution for spot rates is being used, then these rates can be used to compute the bond prices as needed. Otherwise, the tree has to be extended to the maturity date of the longest bond in the basket and bond prices have to be computed using the usual tree methodology. The first solution is faster and less subject to numerical error, but a model with a closed-form solution may or may not be suitable for the problem at hand.

Arbitrage-free pricing models usually assume that some set of securities is fairly priced. In the case of futures, the standard assumption is that the forward prices of all bonds in the deliverable basket are fair. Technically, this calibration can be accomplished by assigning a spread to each bond such that its forward price in the model matches the forward price in the market. The assumption that bond prices are fair is in the spirit of pricing the futures relative to cash. Another approach can be used to decide whether the underlying bonds are fairly priced or not.

Term structure models commonly used for pricing futures fall into two main categories. First are one- or two-factor short-rate models along the lines of those described in Chapters 7–9. These models are relatively easy to implement and, for the most part, flexible enough to capture the yield curve dynamics driving futures prices. With only one or two factors, however, these models cannot capture the idiosyncratic price movements of any deliverable bond relative to another.

The second type of model allows for a richer set of relative price movements across bonds in the deliverable basket. These models essentially allow each bond to follow its own price or yield process. There are costs to this flexibility, however. First, ensuring that these models are arbitrage free takes some effort. Second, the user must specify the parameters that describe the stochastic behavior of all bond prices in the basket, for example, a volatility for each of the bonds in the basket along with correlations between each pair of bonds.

Futures traders often describe their models in terms of the *beta* of each bond in the basket relative to a benchmark bond in the basket. The beta of a bond represents the expected change in yield of that bond given a one-basis-point change in the yield of the benchmark. A bond with a beta of 1.02, for example, implies that the bond's yield is assumed to change 2% more than the yield of the benchmark bond. The beta of a particular bond can be thought of as the coefficient from a regression of changes in its yield on changes in the benchmark's yield. Note that in a one-factor model the beta of a bond is simply the ratio of the volatility of that bond yield to the volatility of the benchmark's yield.

APPENDIX TO CHAPTER 12

Short-Term Rates and Their Derivatives

This appendix builds on Appendixes A11.2 and A11.3 to explain the pricing of forward and futures rates and the futures-forward rate difference in term structure models. All expectations here are also with respect to some risk-neutral or pricing probabilities.

The first step in this Appendix is to show that the futures rate with respect to a term rate, like a Euribor futures rate in the text, is greater than the corresponding forward term rate.

The β-year rate, t years forward, r_t^{fwd}, is the rate that equates the present value at time t of receiving a unit of currency at time $t + \beta$ with the price of a β-year zero coupon bond, t-years forward, $p_0^{fwd}(t)$. Using continuously compounded rates,

$$p_0^{fwd}(t) = e^{-\beta r_t^{fwd}} \tag{A12.1}$$

Furthermore, by Equation (A11.19),

$$p_0^{fwd}(t) = e^{-\beta r_t^{fwd}} > F_0 \tag{A12.2}$$

where F_0 is the futures price of that zero coupon bond.

Let r_t be the realized β-year rate at time t. Then, by definition, the β-year zero coupon bond price at that time, p_t, is,

$$p_t = e^{-\beta r_t} \tag{A12.3}$$

and, by Equation (A11.14),

$$F_0 = E[p_t] = E[e^{-\beta r_t}] \tag{A12.4}$$

Because of daily settlement payments of futures contracts, and following the same argument as in Appendix A11.2, which leads to Equation (A11.14), it can be shown that the futures rate, r_t^{fut}, corresponding to the β-year zero coupon bond rate, is simply today's expectation of that bond rate at time t,

$$r_t^{fut} = E[r_t] \tag{A12.5}$$

By Jensen's inequality and Equation (A12.5),

$$E[e^{-\beta r_t}] > e^{-\beta E[r_t]} = e^{-\beta r_t^{fut}} \tag{A12.6}$$

Stringing together Equations (A12.2), (A12.4), and (A12.6),

$$e^{-\beta r_t^{fwd}} > F_0 = E[e^{-\beta r_t}] > e^{-\beta r_t^{fut}} \tag{A12.7}$$

Finally, focusing on the leftmost and rightmost terms of Equation (A12.7),

$$r_t^{fut} > r_t^{fwd} \tag{A12.8}$$

as was to be proved.

The next step in this appendix is to show that the futures rate with respect to an average rate, like a one-month SOFR (Secured Overnight Financing Rate) futures rate or a fed funds futures rate, is greater than the futures rate with respect to a term rate. This fact, combined with the result just proved, implies that futures rates with respect to average rates are also greater than the corresponding forward rates.

Let $R(t, t+\beta)$ be the integral of the continuously compounded overnight rates from t to $t+\beta$, so that the average rate is $R(t, t+\beta)/\beta$. Once again, because of the daily settlement of futures contracts, today's futures rate on the average rate, A_t^{fut}, is today's expectation of the average at time t,

$$A_t^{fut} = E\left[\frac{1}{\beta}R(t, t+\beta)\right] \tag{A12.9}$$

By definition of the β-year term rate at time t, r_t,

$$e^{-\beta r_t} = E_t[e^{-R(t,t+\beta)}] \tag{A12.10}$$

where $E_t[\cdot]$ gives the expectation as of time t. Applying Jensen's inequality to (A12.10) and rearranging terms,

$$e^{-\beta r_t} > e^{-E_t[R(t,t+\beta)]}$$

$$r_t < \frac{1}{\beta}E_t[R(t, t+\beta)]$$

$$r_t^{fut} < \frac{1}{\beta}E[R(t, t+\beta)] = A_t^{fut} \tag{A12.11}$$

where the third equation takes expectations of the second as of today, and the last equality follows from Equation (A12.9). Equation (A12.11) says that the futures rate on the average is greater than the futures rate on the term rates, as was to be proved.

The final step in this appendix is to show that the futures rate with respect to a compounded rate, like a three-month SOFR futures rate, is greater than the futures rate with respect to an average rate. Once again, because of daily settlement, today's futures rate on a compounded rate, C_t^{fut}, is the expectation of that compounded rate in the future. Mathematically, in a continuously compounded version of Equation (12.10) in the text,

$$C_t^{fut} = \frac{1}{\beta}(E[e^{R(t,t+\beta)}] - 1) \tag{A12.12}$$

By Jensen's inequality and the properties of the exponential function,

$$E[e^{R(t,t+\beta)}] - 1 > e^{E[R(t,t+\beta)]} - 1 > E[R(t, t + \beta)] \tag{A12.13}$$

Finally, combining Equations (A12.12), (A12.13), and the definition of the futures rate on the average in (A12.9),

$$C_t^{fut} = \frac{1}{\beta}(E[e^{R(t,t+\beta)}] - 1) > \frac{1}{\beta}E[R(t, t + \beta)] = A_t^{fut} \tag{A12.14}$$

Hence, as was to be proved, the futures rate on a compounded rate is greater than the futures on an average rate.

To summarize, the financial insight is that, because of daily settlement, the forward price on a zero coupon bond is greater than the futures price on that bond. This result, from Chapter 11, plus a convexity effect in moving from price to rate, shows that the futures rate with respect to a term rate exceeds the corresponding forward term rate. Then, accounting for the convexity effects of moving to a futures rate with respect to an average rate or with respect to a compounded rate, these futures rates exceed the futures rate with respect to the term rate. The bottom line, then, is that all of the futures rates discussed in the chapter exceed the corresponding term forward rate.

APPENDIX TO CHAPTER 13

Interest Rate Swaps

A13.1 PRICING A EURIBOR SWAP AS OF FEBRUARY 24, 2022

This section prices a two-year fixed-for-floating swap, where the floating rate is three-month Euribor. The inputs are the term structure of €STR OIS given in Table A13.1, the two-year Euribor swap rate of 0.078%, and the two-year €STR–Euribor basis swap spread of 0.138%. All cash flows are assumed to follow the actual/360 convention, and all of the relevant dates are assumed to be business days.

The OIS rates in the fourth column of the table are observed in the market. The discount factors are calculated by setting the present value of the fixed payments on each OIS (including the fictional notional amount at maturity) equal to par, which is the value of the floating legs that pay compounded daily €STR. Letting $d(t)$ denote the discount factor for t years, the

TABLE A13.1 €STR OIS Rates as of February 24, 2022.

Term (years)	Term (date)	Term (days)	Rate (%)	Discount Factor	Fwd Rate (%)
0.25	05/24/2022	89	−0.5695	1.0014099	−0.1408
0.50	08/24/2022	181	−0.5580	1.0028134	−0.1400
0.75	11/24/2022	273	−0.5110	1.0038902	−0.1073
1.00	02/24/2023	365	−0.4380	1.0044606	−0.0568
1.25	05/24/2023	454	−0.3380	1.0042784	0.1081
1.50	08/24/2023	546	−0.2330	1.0035455	0.0730
1.75	11/24/2023	638	−0.1400	1.0024888	0.1054
2.00	02/24/2024	730	−0.0600	1.0012201	0.1267

equations determining the discount factors, along the same lines as for SOFR swaps explained in Chapter 2, are the following,

$$
\begin{aligned}
1 &= (1 - 0.5695\% \times 89/360)d(0.25) \\
1 &= (1 - 0.5580\% \times 181/360)d(0.50) \\
1 &= (1 - 0.5110\% \times 273/360)d(0.75) \\
1 &= (1 - 0.4380\% \times 365/360)d(1.00) \\
1 &= -0.3380\% \times 89/360 \times d(0.25) \\
&\quad + (1 - 0.3880\% \times (454 - 89)/360)d(1.25) \\
1 &= -0.2330\% \times 181/360 \times d(0.50) \\
&\quad + (1 - 0.2330\% \times (546 - 181)/360)d(1.50) \\
1 &= -0.1400\% \times 273/360 \times d(0.75) \\
&\quad + (1 - 0.1400\% \times (638 - 273)/360)d(1.75) \\
1 &= -0.0600\% \times 365/360 \times d(1.00) \\
&\quad + (1 - 0.0600\% \times (730 - 365)/360)d(2.00) \qquad \text{(A13.1)}
\end{aligned}
$$

Solving this set of equations gives the discount factors in Table A13.1. The resulting forward rates are shown in the table as well but are not needed for the remaining calculations. In any case, with these discount factors, the two-year Euribor swap rate of 0.078% and the two-year €STR versus three-month Euribor basis swap spread of 0.138% are related by equating the present value of the fixed side of the Euribor swap payments (including the fictional notional amount) to one (i.e., the value of floating €STR, including the fictional notional amount) plus the present value of the basis swap spread payments. Mathematically,

$$
\begin{aligned}
&0.078\% \times 365/360 \times d(1.00) + (1 + 0.078 \times (730 - 365)/360) \times d(2.00) \\
&\quad = 1 + 0.138\% \times 89/360 \times d(0.25) \\
&\qquad + 0.138\% \times (181 - 89)/360 \times d(0.50) + \cdots \\
&\qquad + 0.138\% \times (730 - 638)/360 \times d(2.00) \qquad \text{(A13.2)}
\end{aligned}
$$

Note that Equation (A13.2) is the pricing condition. Given the 0.078% swap rate, it can be used to solve for the 0.138% basis swap spread. Or, conversely, given the 0.138% basis swap spread, it can be used to solve for the 0.078% swap rate.

A13.2 TWO-CURVE PRICING

For the purposes of this section, assume for simplicity that payments are annual at times $t = 1, \ldots, T$. For ease of exposition, say that the risk-free rate index is €STR and that the non-risk-free rate index is Euribor. Given a set of discount factors, $d(t)$, derived from €STR OIS; a set of Euribor swap rates, $c(t)$; and €STR versus Euribor basis swap spreads, $x(t)$, the fair pricing conditions for the Euribor swaps, discussed in the text and the previous section are, for every t,

$$c(t)\sum_{s=1}^{t} d(s) + d(t) = 1 + x(t)\sum_{s=1}^{t} d(s) \tag{A13.3}$$

Now define a set of adjusted Euribor forward rates, $L'(t)$, such that the present values of the floating legs of the Euribor swaps equal their correct values, which are given by the right-hand side of (A13.3),

$$\sum_{s=1}^{t} L'(s)d(s) + (1 + L'(t))d(t) = 1 + x(t)\sum_{s=1}^{t} d(s) \tag{A13.4}$$

Given the basis swap spreads, Equation (A13.4) could be used to solve for all of the adjusted Euribor forward rates, one at a time, starting with $t = 1$ and continuing through to $t = T$. These $L'(t)$ could then be used to price payments that depend on Euribor. However, noticing that the right-hand sides of Equations (A13.3) and (A13.4) are the same, these two equations can be combined,

$$c(t)\sum_{s=1}^{t} d(s) + d(t) = \sum_{s=1}^{t} L'(s)d(s) + (1 + L'(t))d(t) \tag{A13.5}$$

But given all of the swap rates, $c(t)$, Equation (A13.5) can be used iteratively to solve for the $L'(t)$. In other words, so long as the Euribor swaps are priced fairly relative to €STR OIS, there is no need to know the basis swap spreads.

In summary, then, the two-curve methodology for pricing Euribor swaps is as follows. First, solve for the adjusted forward rates, $L'(t)$, as just described. Second, project those adjusted forward rates as future Euribor floating rates to set floating-leg payments. Third, discount both fixed- and floating-leg payments using the €STR discount factors.

APPENDIX TO CHAPTER 14

Corporate Debt and Credit Default Swaps

A14.1 CUMULATIVE DEFAULT AND SURVIVAL RATES

Proposition: If the hazard rate is constant at λ, then the cumulative survival probability to time t, $CS(t)$, is $e^{-\lambda t}$, and the cumulative default probability, $CD(t)$, is $1 - e^{-\lambda t}$.

Proof: The probability of no default to time $t + \Delta t$, $CS(t + \Delta t)$, is the probability that there is no default to time t and no default from then to time $t + \Delta t$. Mathematically,

$$CS(t + \Delta t) = CS(t) \times (1 - \lambda \Delta t) \tag{A14.1}$$

$$\lambda CS(t) = -\frac{CS(t + \Delta t) - CS(t)}{\Delta t} \tag{A14.2}$$

Taking the limit of the right-hand side of (A14.2) as Δt approaches zero is the derivative of $CS(t)$, denoted $CS'(t)$. Therefore,

$$\lambda CS(t) = -CS'(t) \tag{A14.3}$$

$$CS(t) = e^{-\lambda t} \tag{A14.4}$$

where Equation (A14.4) is the solution to Equation (A14.3). The cumulative default probability is just one minus the cumulative survival probability, which gives $CD(t) = 1 - e^{-\lambda t}$.

A14.2 UPFRONT AMOUNTS

Continue with the notation of the previous section, and add the following:

- s^{CDS}: CDS spread
- c^{CDS}: CDS coupon
- T: maturity of the CDS, in years
- n: number of CDS premium payments per year
- t_i: time of CDS premium payment i, $i = 1, \ldots, nT$, with $t_0 \equiv 0$
- $d(t_i)$: discount factor at time t_i, at risk-free or benchmark rates
- R: recovery rate
- UF: upfront amount

Then, following the logic described in the text, with the detail that premium payments follow the actual/360 day-count convention, the expected discounted value of the fee leg, V^{fee}, is,

$$V^{fee} = s^{CDS} \sum_{i=1}^{nT} \frac{t_i - t_{i-1}}{360} CS(t_i)d(t_i) + \frac{1}{2} s^{CDS} \sum_{i=1}^{nT} \frac{t_i - t_{i-1}}{360} [CS(t_{i-1}) - CS(t_i)]d(t_i) \tag{A14.5}$$

The value of the contingent leg, V^{cont}, again following the logic described in the text, is,

$$V^{Cont} = (1 - R) \sum_{i=1}^{nT} [CS(t_{i-1}) - CS(t_i)]d(t_i) \tag{A14.6}$$

A CDS with a premium of s^{CDS} is fair if the hazard rate is such that the values of the fee and contingent legs, as given in Equations (A14.5) and (A14.6), are equal.

The upfront amount, following the logic of the text, equals the expected discount value of paying the CDS coupon as a premium payment rather than the CDS spread. Essentially, the upfront amount is computed like the fee leg, with $-(c^{CDS} - s^{CDS})$ or $(s^{CDS} - c^{CDS})$ replacing s^{CDS}. Hence,

$$UF = (s^{CDS} - c^{CDS}) \sum_{i=1}^{nT} \frac{t_i - t_{i-1}}{360} CS(t_i)d(t_i) + \frac{1}{2}(s^{CDS} - c^{CDS}) \sum_{i=1}^{nT} \frac{t_i - t_{i-1}}{360} [CS(t_{i-1}) - CS(t_i)]d(t_i) \tag{A14.7}$$

A14.3 AN APPROXIMATION FOR CDS SPREADS

For the purposes of this section, assume that premium payments of the CDS are all Δt years apart, that is, $t_i - t_{i-1} = \Delta t$ for all t_i. Substituting that relationship into Equations (A14.5) and (A14.6) and setting the two equations equal, gives the following,

$$s^{CDS}\Delta t \sum_{i=1}^{nT} CS(t_i)d(t_i) + \tfrac{1}{2}s^{CDS}\Delta t \sum_{i=1}^{nT}[CS(t_{i-1}) - CS(t_i)]d(t_i)$$
$$= (1 - R)\sum_{i=1}^{nT}[CS(t_{i-1}) - CS(t_i)]d(t_i) \qquad \text{(A14.8)}$$

Substituting in from Equation (A14.4), each term of (A14.8) can be simplified as follows,

$$\begin{aligned} 0 = {} & s^{CDS}\Delta t e^{-\lambda t_i} \\ & + \frac{1}{2}s^{CDS}\Delta t[e^{-\lambda t_{i-1}} - e^{-\lambda t_i}] \\ & - (1 - R)[e^{-\lambda t_{i-1}} - e^{-\lambda t_i}] \end{aligned} \qquad \text{(A14.9)}$$

Furthermore, because $t_i - t_{i-1} = \Delta t$ for all t_i, it is easy to show that, if Equation (A14.9) holds for t_i, then it also holds for t_{i+1}. Hence, in this special case, s^{CDS} can be solved from any one date. Proceeding then from Equation (A14.9), multiply through by $e^{-\lambda t_i}$ and simplify to see that,

$$s^{CDS}\Delta t\left[1 + \frac{1}{2}(e^{\lambda\Delta t} - 1)\right] = (1 - R)[e^{\lambda\Delta t} - 1] \qquad \text{(A14.10)}$$

$$s^{CDS}\Delta t\frac{[1 + e^{\lambda\Delta t}]}{2} = (1 - R)[e^{\lambda\Delta t} - 1] \qquad \text{(A14.11)}$$

$$s^{CDS} = (1 - R)\left[\frac{2}{\Delta t}\frac{e^{\lambda\Delta t} - 1}{e^{\lambda\Delta t} + 1}\right] \qquad \text{(A14.12)}$$

Finally, take the limit as Δt approaches zero of the terms in the hard bracket on the right-hand side of (A14.12) to obtain,

$$s^{CDS} \approx \lambda(1 - R) \qquad \text{(A14.13)}$$

A14.4 CDS-EQUIVALENT BOND SPREADS

Following the logic of the text and this appendix, the expected discounted value of a bond's coupons can be computed along the lines of the fee leg of a CDS. Similarly, the expected discounted value of a bond's principal payment can be computed along the lines of the contingent leg, except that the payment from the bond upon default is R times the principal amount, while the payment due to the buyer of protection upon default is $1 - R$ times the principal amount. Therefore, using Equations (A14.5) and (A14.6) – simplified here to assume that coupon payments are exactly one-half years apart – the price of a bond, P, with a coupon rate, c, given a hazard rate, is given by,

$$
\begin{aligned}
P &= \frac{c}{2}\sum_{i=1}^{nT} CS(t_i)d(t_i) \\
&+ \frac{c}{2}\sum_{i=1}^{nT}[CS(t_{i-1}) - CS(t_i)]d(t_i)] \\
&+ R\sum_{i=1}^{nT}[CS(t_{i-1}) - CS(t_i)]d(t_i)
\end{aligned}
\quad \text{(A14.14)}
$$

To solve for the CDS-equivalent bond spread, find the hazard rate that solves Equation (A14.14). Then, using that hazard rate, find the s^{CDS} that sets the fee leg in Equation (A14.5) equal to the contingent leg in Equation (A14.6).

A14.5 BOND SPREAD WITH MARKET RECOVERY

From the definition of bond spread, if the bond does not default (and rates do not change), then the return on the bond equals the risk-free or benchmark rate, r, plus the spread. If the bond defaults and recovers R^m of its market price, then the return over the moment of default is $(R^m P - P)/P$ or $(R^m - 1)$. Therefore, over a short time interval dt years, over which the probability of no default and default are $1 - \lambda dt$ and λdt, respectively, the expected return on the bond is,

$$
\begin{aligned}
&(1 - \lambda dt) \times (r + s)dt + \lambda dt \times (R^m - 1) \\
&\approx (r + s)dt + \lambda dt(R^m - 1)
\end{aligned}
\quad \text{(A14.15)}
$$

where the approximation follows from ignoring the very small terms, that is, those that are multiplied by $(dt)^2$.

Assuming that investors are risk neutral or that the hazard rate is a risk-neutral pricing rate, investors are indifferent between buying a corporate bond and buying a bond without default risk if the expected return of the former, given in Equation (A14.15), is equal to the risk-free rate. Mathematically, then,

$$
\begin{aligned}
rdt &= (r + s)dt + \lambda dt(R^m - 1) \\
s &= \lambda(1 - R^m)
\end{aligned}
\quad \text{(A14.16)}
$$

APPENDIX TO CHAPTER 15

Mortgages and Mortgage-Backed Securities

A15.1 MONTH-END BALANCES

This section shows that, under the principal amortization schedule described in the text, the balance outstanding at any time equals the present value of the remaining payments at the original mortgage rate. Let N be the term of the mortgage, in months; let r be the mortgage rate; let X be the monthly payment, and let $B(i)$ be the balance outstanding at the end of month i, $i = 0, 1, \ldots N$. By definition, as discussed in the text,

$$B(0) = X\frac{12}{r}\left[1 - \left(1 + \frac{r}{12}\right)^{-N}\right] \tag{A15.1}$$

If the balance outstanding at the end of month i does equal the present value of the remaining payments at the rate r, then,

$$B(i) = X\frac{12}{r}\left[1 - \left(1 + \frac{r}{12}\right)^{i-N}\right] \tag{A15.2}$$

$$B(i+1) = X\frac{12}{r}\left[1 - \left(1 + \frac{r}{12}\right)^{i+1-N}\right] \tag{A15.3}$$

According to the logic of the amortization table, the interest component of the payment for month $i+1$ is $(r/12)B(i)$, and the principal component is $X - (r/12)B(i)$. Because $B(0)$ is, by definition, the present value of the remaining payments at the start of the mortgage, this section needs to prove that, for any $i > 0$,

$$B(i) - B(i+1) = X - \frac{r}{12}B(i) \tag{A15.4}$$

To prove this, rearrange terms and then substitute for $B(i)$ and $B(i+1)$ from Equations (A15.2) and (A15.3), respectively,

$$\left(1+\frac{r}{12}\right)B(i) - B(i+1) \stackrel{?}{=} X \qquad \text{(A15.5)}$$

$$X\frac{12}{r}\left[1+\frac{r}{12}-\left(1+\frac{r}{12}\right)^{i+1-N}-1+\left(1+\frac{r}{12}\right)^{i+1-N}\right] \stackrel{?}{=} X \qquad \text{(A15.6)}$$

$$X\frac{12}{r}\left[\frac{r}{12}\right] \stackrel{?}{=} X \qquad \text{(A15.7)}$$

$$X = X \qquad \text{(A15.8)}$$

which is clearly true for any parameters of the problem.

A15.2 PRICING MBS WITH TERM STRUCTURE MODELS

This section very briefly discusses three problems that arise in the context of pricing MBS in term structure models: *path dependence*, proxies for the mortgage rate, and factors apart from interest rates.

As for path dependence, the burnout and media effects imply that the value of an MBS depends not only on the current term structure of interest rates, but also on the history of the term structure. Pricing by backward induction, however, described in earlier chapters of the book, does not naturally accommodate path dependence: at any node of a binomial tree, for example, there is no memory of how the interest rate process moved from its current state to that particular node of the tree.

A popular solution for pricing path-dependent claims is *Monte Carlo simulation*. To price a security in this framework, in a one-factor setting, proceed as follows. First, generate a large number of paths of interest rates at the frequency and to the horizon desired. For this purpose, paths are generated using a particular risk-neutral process for the short-term rate. Second, calculate the cash flows of the security along each path. In the mortgage context, this would include the security's scheduled payments along with its prepayments. Burnout and the media effect can be implemented in this framework, because each path is available in its entirety as cash flows are calculated. Third, starting at the end of each path, calculate the discounted value of the security's cash flows along each path using the interest rates along that path. Fourth, compute the value of the security as the average of the discounted values across paths.

To connect Monte Carlo simulation with the pricing approach used elsewhere in the book, recall Equation (A11.5), reproduced here for convenience,

$$P_0 = E\left[\frac{P_n}{\prod_{i=0}^{n-1}(1+r_i)}\right] \tag{A15.9}$$

where r_i is the short-term rate in period i, P_n is the value of a claim in n periods, and P_0 is the price of the claim today. In light of the discussion in this section, the term inside the hard brackets is analogous to the discounted value of a security along one path. The expectation is analogous to the average of those discounted values across paths.

In the Monte Carlo framework, measures of interest rate sensitivity can be computed by shifting the initial term structure in some manner, repeating the valuation process, and calculating the difference between the initial and shifted prices. As a numerical matter, paths should not be regenerated between the initial and shifted valuations, as that would introduce noise into the sensitivity calculations.

A general drawback of Monte Carlo simulation is that is does not naturally accommodate the valuation of American- or Bermudan-style options, because the value of holding an option at any state typically requires pricing by backward induction. While methodological advances are now used to overcome this problem,[1] the issue does not really arise when valuing the prepayment option. As it is generally accepted that homeowner behavior is not well explained as the optimal exercise of a fixed income option, a path-dependent prepayment function, which is well-suited to Monte Carlo simulation, is the preferred approach.

The second problem addressed in this section relates to the mortgage rate. Valuing an MBS along a path requires both the benchmark of discounting rates as well as the mortgage rate. Discounting can be done at benchmark rates plus a spread, but the incentive of a prepayment model depends on the current mortgage rate. The difficulty is that calculating the fair mortgage rate at a single date on a single path of a Monte Carlo simulation is a problem of the same order of magnitude as the original problem of pricing a particular MBS as of the current state! Common practice, therefore, is to build a simple model of the mortgage rate as a function of benchmark rates and, perhaps, volatility, for example, an estimated regression model of the mortgage rate as a function of one or more benchmark rates and rate volatility. It may not be trivial to compute long-term benchmark rates along a path of short-term

[1]Longstaff, F., and Schwartz, E. (2001), "Valuing American Options by Simulation: A Simple Least-Squares Approach," *Review of Financial Studies* 14(1), Spring, pp. 113-147.

rates, but this difficulty can usually be overcome with a closed-form solution for long-term rates or with a numerical approximation consistent with the short-term rate process.

The third problem is that defaults and turnover depend on variables other than interest rates, like housing prices, which are not traditionally included in the pricing of fixed income securities. Incorporating these variables, including their correlations with interest rates, is yet another challenge of pricing MBS.

APPENDIX TO CHAPTER 16

Fixed Income Options

A16.1 THEORETICAL FOUNDATIONS FOR APPLYING BLACK-SCHOLES-MERTON (BSM) TO SELECTED FIXED INCOME OPTIONS

The justification for applying BSM in each of the cases of the text takes the following form:

1. Given the functional form of a probability distribution (e.g., normal, lognormal), there exist parameters of that distribution such that V_0, the arbitrage-free price of any asset today, is given by,

$$\frac{V_0}{N_0} = E_0\left[\frac{V_t}{N_t}\right] \tag{A16.1}$$

   where N_t is the price at time t of an asset chosen as the *numeraire*; V_t is value at time t of an asset being priced today, including reinvested cash flows; and $E_t[\cdot]$ gives expectations as of time t under the appropriately parameterized probability distribution. Equation (A16.1) is known as the *martingale property* of asset prices. This claim is proved in a special case in Section A16.2 but used more generally here.
2. Say that the rate or security price underlying an option at time t is S_t. It follows from the previous point that the value of a call option with strike K and time to expiry T is,

$$V_0^{Call} = N_0 E_0\left[\frac{(S_T - K)^+}{N_T}\right] \tag{A16.2}$$

   while the value of a put is,

$$V_0^{Put} = N_0 E_0\left[\frac{(K - S_T)^+}{N_T}\right] \tag{A16.3}$$

3. In the contexts of the text, it is possible to choose the numeraire such that,

$$S_0 = E_0[S_T] \tag{A16.4}$$

and such that Equations (A16.2) and (A16.3) can be written as, respectively,

$$V_0^{Call} = h_0 E_0[(S_T - K)^+] \tag{A16.5}$$

$$V_0^{Put} = h_0 E_0[(K - S_T)^+] \tag{A16.6}$$

for some h_0 that is known as of time 0. This is proven in Section A16.3.

4. If S_t has a normal distribution with volatility parameter σ, then Section A16.4 shows that (A16.4) through (A16.6) become the normal BSM-style formulae,

$$V_0^{Call} = h_0 \xi^N(S_0, T, K, \sigma) \tag{A16.7}$$

$$V_0^{Put} = h_0 \pi^N(S_0, T, K, \sigma) \tag{A16.8}$$

for the functions $\xi^N(\cdot)$ and $\pi^N(\cdot)$ defined in that section. On the other hand, if S_t has a lognormal distribution with volatility parameter σ, Section A16.4 shows that (A16.4) through (A16.6) become the lognormal BSM-style formulae,

$$V_0^{Call} = h_0 \xi^{LN}(S_0, T, K, \sigma) \tag{A16.9}$$

and

$$V_0^{Put} = h_0 \pi^{LN}(S_0, T, K, \sigma) \tag{A16.10}$$

for the functions $\xi^{LN}(\cdot)$ and $\pi^{LN}(\cdot)$ defined in that section.

A16.2 NUMERAIRES, PRICING MEASURES, AND THE MARTINGALE PROPERTY

Define the following:

- *Gains process*. The gains process of an asset at any time equals the value of that asset at that time plus the value of all its cash flows reinvested to that time. For this purpose, all cash flows are reinvested and then rolled at prevailing short-term interest rates.

- *Numeraire asset*. Given a particular numeraire asset, the gains process of any other asset can be expressed in terms of the numeraire asset by dividing that gains process by the gains process of the numeraire asset. The gains process of a security in terms of the numeraire asset is called the *normalized gains process* of that asset.

For concreteness, Table A16.1 gives an example of these concepts. The asset under consideration is a long-term, 4% coupon bond with a face amount of 100. The numeraire is a two-year zero coupon bond with a unit face amount. The gains process is observed today, after one year, and after two years.

The realization of the short-term rate, which in this example is the one-year rate, is given in row (i). The realization of the bond price over time is given in row (ii). A 4% coupon on 100 is paid on dates 1 and 2 and shown in row (iii) and row (iv). The payment on date 1 is reinvested for one year at the short-term rate on date 1, that is, 2%. The gains process of the bond given in row (v) is the sum of its price and reinvested cash flows, i.e., the sums of rows (ii) through (iv). The price realization of the two-year zero coupon bond, which is the chose numeraire, is given in row (vi). Finally, the normalized bond gains process, given in row (vii), is the bond gains process divided by the price of the numeraire, i.e., row (v) divided by row (vi).

These definitions allow for the statement of the main result of this section: in the absence of arbitrage opportunities, there exists a parameterization of a given probability distribution, or a *pricing measure*, such that the normalized gains of any asset today equals the expected value of that asset's normalized gains in the future. Technically, there exist probabilities such that the normalized gains process is a *martingale*. As the goal here is intuition rather than mathematical generality, this result is proven in the context of a single-period, binomial process.

TABLE A16.1 Example of the Calculation of a Normalized Gains Process

		End-of-Year Realizations		
		0	**1**	**2**
(i)	Short-Term/1-Year Rate	1%	2%	1.5%
(ii)	Bond Price	100	95	97.50
(iii)	Date 1 Reinvested Coupon		4	4(1.02)=4.08
(iv)	Date 2 Reinvested Coupon			4
(v)	Gains Process	100	99	105.58
(vi)	Price of 2-Year Zero/Numeraire	0.9612	0.9804	1.0
(vii)	Normalized Bond Gains Process	104.04	100.98	105.58

The starting point is state 0 of date 0, after which the economy moves to either state 0 or state 1 of date 1. Three assets will be considered, A, B, and C, with current prices A_0, B_0, and C_0, and date 1, state i prices of A_1^i, B_1^i, and C_1^i. Without loss of generality here, the date 1 prices include any cash flows of the securities on date 1.

In this framework, any asset can be priced by arbitrage relative to the other two assets. The method is just as in Chapter 7. To price asset C by arbitrage, construct its replicating portfolio, in particular, a portfolio with α of asset A and β of asset B such that,

$$C_1^0 = \alpha A_1^0 + \beta B_1^0 \tag{A16.11}$$

$$C_1^1 = \alpha A_1^1 + \beta B_1^1 \tag{A16.12}$$

Then, to rule out risk-free arbitrage opportunities, it must be the case that,

$$C_0 = \alpha A_0 + \beta B_0 \tag{A16.13}$$

Now let asset A be the numeraire and rewrite equations (A16.11) through (A16.13) in terms of the normalized gains processes of assets B and C. To do this, simply divide each of the equations by the corresponding value of the numeraire asset A, that is, divide (A16.11) by A_1^0, (A16.12) by A_1^1, and (A16.13) by A_0. Furthermore, denote the normalized gains process of the assets by $\overline{B}$ and $\overline{C}$. Then, equations (A16.11) through (A16.13) become,

$$\overline{C}_1^0 = \alpha + \beta \overline{B}_1^0 \tag{A16.14}$$

$$\overline{C}_1^1 = \alpha + \beta \overline{B}_1^1 \tag{A16.15}$$

$$\overline{C}_0 = \alpha + \beta \overline{B}_0 \tag{A16.16}$$

Furthermore, solving (A16.14) and (A16.15) for α and β,

$$\alpha = \frac{\overline{B}_1^1 \overline{C}_1^0 - \overline{B}_1^0 \overline{C}_1^1}{\overline{B}_1^1 - \overline{B}_1^0} \tag{A16.17}$$

$$\beta = \frac{\overline{C}_1^1 - \overline{C}_1^0}{\overline{B}_1^1 - \overline{B}_1^0} \tag{A16.18}$$

In the framework just described, it is now shown that there exists a pricing measure such that the expected normalized gains process of each security is a martingale. More specifically, there is a probability p of moving to state 1 of date 1 (and $1-p$ of moving to state 0 of date 1) such that the expected value of the normalized gain of each security on date 1 equals its normalized gain on date 0. Mathematically, it has to be shown that there is a p such that,

$$\overline{C}_0 = p\overline{C}_1^1 + (1-p)\overline{C}_1^0 \tag{A16.19}$$

$$\overline{B}_0 = p\overline{B}_1^1 + (1-p)\overline{B}_1^0 \tag{A16.20}$$

Solving (A16.20) for p gives,

$$p = \frac{\overline{B}_0 - \overline{B}_1^0}{\overline{B}_1^1 - \overline{B}_1^0} \tag{A16.21}$$

But this value of p also solves (A16.19). To see this, start by substituting p from (A16.21) into the right-hand side of (A16.19),

$$p\overline{C}_1^1 + (1-p)\overline{C}_1^0 = \frac{\overline{B}_0 - \overline{B}_1^0}{\overline{B}_1^1 - \overline{B}_1^0}\overline{C}_1^1 - \frac{\overline{B}_0 - \overline{B}_1^1}{\overline{B}_1^1 - \overline{B}_1^0}\overline{C}_1^0 \tag{A16.22}$$

$$= \overline{B}_0 \frac{\overline{C}_1^1 - \overline{C}_1^0}{\overline{B}_1^1 - \overline{B}_1^0} + \frac{\overline{B}_1^1\overline{C}_1^0 - \overline{B}_1^0\overline{C}_1^1}{\overline{B}_1^1 - \overline{B}_1^0} \tag{A16.23}$$

$$= \overline{B}_0\beta + \alpha \tag{A16.24}$$

$$= \overline{C}_0 \tag{A16.25}$$

Equation (A16.23) just rearranges the terms of (A16.22); combining (A16.23) with (A16.17) and (A16.18) gives (A16.24); and (A16.24) with (A16.16) gives (A16.25). Hence, as was to be shown, there is a pricing measure, in this case the probability p, such that the normalized gains processes of B and C are martingales. And, of course, since nothing distinguishes A from the other assets, a probability with the same properties could have been found had B or C been chosen as the numeraire instead.

A16.3 CHOOSING THE NUMERAIRE AND BSM PRICING

In the contexts of this chapter, it is possible to choose a numeraire such that the underlying is a martingale and such that the value of a call is given by $h_0\xi^N(S_0, T, K, \sigma)$ or $h_0\xi^{LN}(S_0, T, K, \sigma)$, in the normal or lognormal cases, respectively, and the value of a put by $h_0\pi^N(S_0, T, K, \sigma)$ or $h_0\pi^{LN}(S_0, T, K, \sigma)$ in the normal or lognormal cases, where those functions are defined in Section A16.4. This section gives the appropriate definition of the underlying, the appropriate numeraire, and the resulting quantity h_0.

A16.3.1 Bond Options

Start with a European-style option, expiring on date T, written on a longer-term bond. The underlying of this option is a forward position in the bond for delivery on date T. It is first shown that taking the zero coupon bond maturing at time T to be the numeraire makes this forward bond price a martingale. Proving this is somewhat complex, because the gains process of a bond includes reinvested coupons. Therefore, to keep the presentation simple, the martingale result is derived in a three-date, two-period setting. The current date is date 0, and the expiration or forward delivery date is date 2. The bond is assumed to pay a coupon c on each of dates 1 and 2, and its price at time t is denoted B_t. The numeraire is the zero coupon bond maturing on date 2 with a price, on date t, of $d_t(2)$. Lastly, let r_1 denote the current one-period rate; r_2 the one-period rate, realized one period from now; and f the current one-period rate, one-period forward.

Under these assumptions, and the expression of the zero coupon bond price on various dates in terms of the prevailing one-period rates, the gains process of the bond on the three dates is given by the expressions,

- Date 0: $\frac{B_0}{d_0(2)} = B_0(1+r_1)(1+f)$;
- Date 1: $\frac{B_1+c}{d_1(2)} = (B_1+c)(1+r_2)$;
- Date 2: $\frac{B_2+c(1+r_2)+c}{d_2(2)} = B_2 + c(1+r_2) + c$

Therefore, the martingale property for the bond says that,

$$\frac{B_0}{d_0(2)} = E_0\left[\frac{B_2 + c(1+r_2) + c}{d_2(2)}\right] \qquad \text{(A16.26)}$$
$$B_0(1+r_1)(1+f) = E_0[B_2 + c(1+r_2) + c]$$

The term $c(1+r_2)$ in the expectation on the right-hand side of (A16.26) requires some attention, because r_2 is not known as of date 0. The date-0 value of a payment of $c(1+r_2)$ on date 2 is, however, by the definition of forward rates,

$$\frac{c(1+f)}{(1+r_1)(1+f)} = \frac{c}{1+r_1} \qquad \text{(A16.27)}$$

So, applying the martingale property under the numeraire to a payment of $c(1 + r_2)$ on date 2 requires that,

$$\begin{aligned} \frac{\frac{c}{1+r_1}}{d_0(2)} &= E_0\left[\frac{c(1+r_2)}{d_2(2)}\right] \\ c(1+f) &= E_0[c(1+r_2)] \end{aligned} \tag{A16.28}$$

With this result, the discussion returns to the martingale property of the bond in (A16.26). Substituting (A16.28) into (A16.26),

$$\begin{aligned} B_0(1+r_1)(1+f) - c(1+f) - c &= E_0[B_2] \\ \left[B_0 - \frac{c}{1+r_1} - \frac{c}{(1+r_1)(1+f)}\right](1+r_1)(1+f) &= E_0[B_2] \\ B_0(2) &= E_0[B_2] \end{aligned} \tag{A16.29}$$

The left-hand side of the second line of (A16.29) is the date 0 forward price of the bond for delivery on date 2. The third line, then, simply denotes this forward price by $B_0(2)$. Hence, taking the zero coupon bond of maturity T as a numeraire, the forward price of a bond for delivery on date T is a martingale.

Turning now to the price of an option on the bond, consider a call with payoff $(B_T - K)^+$. Applying the martingale property to the option price and assuming that the forward bond price is lognormal with volatility parameter σ, the call option is priced as,

$$\frac{V_0^{BondCall}}{d_0(T)} = E_0\left[\frac{(B_T - K)^+}{d_T(T)}\right] \tag{A16.30}$$

$$= E_0[(B_T - K)^+] \tag{A16.31}$$

$$V_0^{BondCall} = d_0(T)\xi^{LN}(B_0(T), T, K, \sigma) \tag{A16.32}$$

An analogous argument for a put shows that,

$$V_0^{BondPut} = d_0(T)\pi^{LN}(B_0(T), T, K, \sigma) \tag{A16.33}$$

A16.3.2 Euribor Futures Options

The terminal payoff of a Euribor futures call option with strike K and expiration time T is, per unit notional,

$$[K - f_T(T, T + \tau)]^+ \tag{A16.34}$$

Given the daily settlement feature of Euribor futures options, the numeraire of choice is the money market account, the value of one unit of currency invested and then rolled every period, at the prevailing short-term rate. Denoting the money market account by $M(t)$ and the short-term rate from time $t - 1$ to t by r_t,

$$M(0) = 1 \tag{A16.35}$$

$$M(T) = (1 + r_1)(1 + r_2) \cdots (1 + r_T) \tag{A16.36}$$

The first point to make about the money market account is that it is the numeraire of the risk-neutral short-term rate process used in the term structure models presented earlier in the book. To see this, apply the martingale property with the numeraire to an arbitrary gains process V_t at time t,

$$\begin{aligned} \frac{V_0}{M(0)} &= E_0\left[\frac{V_T}{M(T)}\right] \\ V_0 &= E_0\left[\frac{V_T}{(1 + r_1)(1 + r_2) \cdots (1 + r_T)}\right] \end{aligned} \tag{A16.37}$$

But the second line of (A16.37) is just the condition that the value of a claim today equals its expected discounted value.

The second point to make about the money markets as numeraire is that futures prices are martingales under this numeraire. This is proved in Section A16.5.

Turning now to Euribor futures options, because they are subject to daily settlement and are futures contracts, their prices are also martingales with the money market account as numeraire. Furthermore, if F_t is the underlying futures price at time t, then at the expiration of a put option on the futures price (call on rates) at time T, the option is worth $(F_T - K)^+$. Putting together the martingale property of the futures, (A16.38), the martingale property of futures options, (A16.39), and the final settlement price of the futures options, (A16.40), results in the price of the Euribor futures put option at time t, denoted V_t^{EBPut},

$$F_0 = E[F_T] \tag{A16.38}$$

$$V_0^{EBPut} = E_0[V_T^{EBPut}] \tag{A16.39}$$

$$= E_0[(F_T - K)^+] \tag{A16.40}$$

Assuming now that F_T is normally distributed, applying Section A16.4 to Equations (A16.38) and (A16.40) shows that,

$$V_0^{EBPut} = \xi^N(F_0, T, K, \sigma) \tag{A16.41}$$

Similarly, for the Euribor futures call option (put on rates),

$$V_t^{EBCall} = \pi^N(F_0, T, K, \sigma) \tag{A16.42}$$

A16.3.3 Bond Futures Options

As shown in Section A16.5, futures prices are a martingale in the risk-neutral measure, that is, when the numeraire is the money market account, $M(t)$. Hence, with F_t the underlying bond futures price at time t,

$$F_0 = E[F_T] \tag{A16.43}$$

By the martingale property, the price of a put option on the futures is,

$$\frac{V_0^{FutPut}}{M(0)} = E_0\left[\frac{(K - F_T)^+}{M(T)}\right] \tag{A16.44}$$

Then, by the definition of the money market account,

$$V_0^{FutPut} = E_0\left[\frac{(K - F_T)^+}{(1 + r_1)(1 + r_2)\cdots(1 + r_T)}\right] \tag{A16.45}$$

To continue, make the assumption – defended in the text – that the discount factor is uncorrelated with the futures price. Then, Equation (A16.45) becomes,

$$V_0^{FutPut} = E_0\left[\frac{1}{(1 + r_1)(1 + r_2)\cdots(1 + r_T)}\right] E_0[(K - F_T)^+] \tag{A16.46}$$

$$= d_0(T)E_0[(K - F_T)^+] \tag{A16.47}$$

where (A16.47) follows from the risk-neutral pricing of a zero coupon bond.

Finally, applying Section A16.5 to (A16.43), (A16.47) with the assumption that the bond futures price has a lognormal distribution,

$$V_0^{FutPut} = d_0(T)\pi^{LN}(F_0, T, K, \sigma) \tag{A16.48}$$

For calls, the analogous result is,

$$V_0^{FutCall} = d_0(T)\xi^{LN}(F_0, T, K, \sigma) \tag{A16.49}$$

A16.3.4 Caplets

Caplets that mature at time T are written on a forward rate from time T to $T+\tau$, whose value, at time t, is denoted by $f_t(T, T+\tau)$. It is first shown that taking a $T+\tau$-year zero coupon bond as the numeraire makes this forward rate a martingale. Let $d_t(T)$ be the time-t price of a zero coupon bond maturing at time T. By the definition of a forward rate of term τ,

$$\begin{aligned} f_t(T, T+\tau) &= \frac{1}{\tau}\left(\frac{d_t(T)}{d_t(T+\tau)} - 1\right) \\ &= \frac{1}{\tau}\left(\frac{d_t(T) - d_t(T+\tau)}{d_t(T+\tau)}\right) \end{aligned} \tag{A16.50}$$

Next, consider a portfolio that is long a T-year zero and short a $T+\tau$-year zero. Taking the $T+\tau$-year zero as the numeraire, the normalized gains process of this portfolio is a martingale. Mathematically,

$$\begin{aligned} \frac{1}{\tau}\left(\frac{d_t(T) - d_t(T+\tau)}{d_t(T+\tau)}\right) &= \frac{1}{\tau}E_t\left[\frac{d_T(T) - d_T(T+\tau)}{d_T(T+\tau)}\right] \\ &= \frac{1}{\tau}E_t\left[\frac{d_T(T)}{d_T(T+\tau)} - 1\right] \\ &= E_t[f_T(T, T+\tau)] \end{aligned} \tag{A16.51}$$

where the last line of (A16.51) just uses the definition of the forward rate. Combining (A16.50) and (A16.51) shows that the forward rate is a martingale under the chosen numeraire,

$$f_t(T, T+\tau) = E_t[f_T(T, T+\tau)] \tag{A16.52}$$

Turning to the valuation of the caplet, its normalized gains process is a martingale as well. Hence, taking expectations of its normalized gain as of $T+\tau$,

$$\frac{V_0^{Caplet}}{d_0(T+\tau)} = E_0\left[\frac{\tau(f_T(T, T+\tau) - K)^+}{d_{T+\tau}(T+\tau)}\right] \tag{A16.53}$$

$$= E_0[\tau(f_T(T, T+\tau) - K)^+] \tag{A16.54}$$

Finally, assuming that the forward rate $f_T(T, T+\tau)$ is normal with variance $\sigma^2 T$, and knowing from (A16.52) with $t = 0$ that its mean is $f_0(T, T+\tau)$, the results of Section A16.4 apply and,

$$V_0^{Caplet} = d_0(T+\tau)\tau\xi^N(f_0(T, T+\tau), T, K, \sigma) \tag{A16.55}$$

A16.3.5 Swaptions

The underlying of a T-year into τ-year swaption is the forward par swap rate from T to $T+\tau$, which, at time t, is denoted by $C_t(T, T+\tau)$. It is first shown that taking an annuity from T to $T+\tau$ as the numeraire makes this forward par swap rate a martingale. Denote the price of this annuity by $A_t(T, T+\tau)$.

Consider receiving the fixed-rate K on a swap from T to $T+\tau$. Its value at time t is,

$$[K - C_t(T, T+\tau)]A_t(T, T+\tau) \tag{A16.56}$$

Applying the martingale property with this annuity as numeraire,

$$\frac{[K - C_t(T, T+\tau)]A_t(T, T+\tau)}{A_t(T, T+\tau)} = E_t\left[\frac{[K - C_T(T, T+\tau)]A_T(T, T+\tau)}{A_T(T, T+\tau)}\right]$$
$$C_t(T, T+\tau) = E_t[C_T(T, T+\tau)] \tag{A16.57}$$

Hence, as claimed, the forward par swap rate is a martingale under this numeraire.

To price a receiver swaption, note that the payoff is $[K - C_T(T, T+\tau)]^+ \times A_T(T, T+\tau)$. Therefore, its value can be calculated as the expectation of its normalized payoff using the same numeraire,

$$\begin{aligned} \frac{V_0^{Receiver}}{A_0(T, T+\tau)} &= E_0\left[\frac{(K - C_T(T, T+\tau))^+ A_T(T, T+\tau)}{A_T(T, T+\tau)}\right] \\ &= E_0[(K - C_T(T, T+\tau))^+] \\ V_0^{Receiver} &= A_0(T, T+\tau)\pi^N(C_0(T, T+\tau), T, K, \sigma) \end{aligned} \tag{A16.58}$$

The last line of (A16.58) follows from (A16.57), the assumption that the forward par swap rate is normal with variance $\sigma^2 T$, and the appropriate result from Section A16.4.

Similarly, a payer option under the assumption of normality has the value,

$$V_0^{Payer} = A_0(T, T+\tau)\xi^N(C_0(T, T+\tau), T, K, \sigma) \tag{A16.59}$$

A16.4 EXPECTATIONS FOR BLACK-SCHOLES-MERTON STYLE OPTION PRICING

As the results in this section are part of the option pricing literature, they are presented here for easy reference but without proof. Let $E^N[\cdot]$ and $E^{LN}[\cdot]$

denote the expectations operators under the normal and lognormal distributions, respectively, and let $N(\cdot)$ denote the standard normal cumulative distribution.

If S_T is normally distributed with means S_0 and variance $\sigma^2 T$, then,

$$\begin{aligned}\xi^N(S_0, T, K, \sigma) &\equiv E_0^N[(S_T - K)^+] \\ &= (S_0 - K)N(d) + \frac{\sigma\sqrt{T}}{\sqrt{2\pi}}e^{-\frac{1}{2}d^2}\end{aligned} \tag{A16.60}$$

$$\begin{aligned}\pi^N(S_0, T, K, \sigma) &\equiv E_0^N[(K - S_T)^+] \\ &= (K - S_0)N(-d) + \frac{\sigma\sqrt{T}}{\sqrt{2\pi}}e^{-\frac{1}{2}d^2}\end{aligned} \tag{A16.61}$$

$$d = \frac{S_0 - K}{\sigma\sqrt{T}} \tag{A16.62}$$

If S_T is lognormally distributed with mean S_0 and variance $S_0^2(e^{\sigma^2 T} - 1)$, then,

$$\begin{aligned}\xi^{LN}(S_0, T, K, \sigma) &\equiv E_0^{LN}[(S_T - K)^+] \\ &= S_0 N(d_1) - KN(d_2)\end{aligned} \tag{A16.63}$$

$$\begin{aligned}\pi^{LN}(S_0, T, K, \sigma) &\equiv E_0^{LN}[(K - S_T)^+] \\ &= KN(-d_2) - S_0 N(-d_1)\end{aligned} \tag{A16.64}$$

$$d_1 = \frac{\ln(\frac{S_0}{K}) + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}} \tag{A16.65}$$

$$d_2 = d_1 - \sigma\sqrt{T} \tag{A16.66}$$

A16.5 FUTURES PRICES ARE MARTINGALES WITH THE MONEY MARKET ACCOUNT AS A NUMERAIRE

The initial value of a futures contract is zero; subsequent cash flows are from daily settlements; and at maturity, the futures price is determined by some final settlement rule. Consider a two-period, three-date framework for simplicity, and let the futures price on date t be F_t. Then, the normalized gains process is,

- Date 0: $\frac{V_0}{M(0)} = 0$;

- Date 1: $\frac{V_1}{M(1)} = \frac{F_1 - F_0}{1+r_1}$;
- Date 2: $\frac{V_2}{M(2)} = \frac{(F_1 - F_0)(1+r_2) + F_2 - F_1}{(1+r_1)(1+r_2)}$

Since the value of a futures contact on date 0 is zero, the martingale property implies that the expectation of the normalized gains at any future date is zero. In particular, for date 1,

$$0 = E_0 \left[\frac{F_1 - F_0}{1 + r_1} \right] \tag{A16.67}$$

But since r_1 is known as of date 0, it follows from (A16.67) that,

$$F_0 = E_0[F_1] \tag{A16.68}$$

As of date 2, the martingale property says that

$$0 = E_0 \left[\frac{(F_1 - F_0)(1 + r_2) + F_2 - F_1}{(1 + r_1)(1 + r_2)} \right] \tag{A16.69}$$

Using the law of iterated expectations, and the fact that r_1 is known as of date 0,

$$0 = E_0 \left[\frac{1}{1 + r_2} E_1[F_2 - F_1] \right] \tag{A16.70}$$

But, since F_1 is known as of date 1, (A16.70) implies that,

$$F_1 = E_1[F_2] \tag{A16.71}$$

Finally then, combine (A16.68) and (A16.71) to see that,

$$F_0 = E_0[E_1[F_2]] = E_0[F_2] \tag{A16.72}$$

Together with (A16.68), (A16.72) shows that the futures price is a martingale under the money-market account or risk-neutral measure, as desired.

About the Website

This book is accompanied by a companion website for instructors: www.wiley.com\go\tuckman\fixedincomesecurities4e

The website includes:

- PowerPoint and PDF of figures and tables that appear in the text, with which instructors can create their own lecture slides
- A sample syllabus
- Sample problems and solutions
- Spreadsheets with calculations supporting sample problems, which can also be used to create additional problems

Index