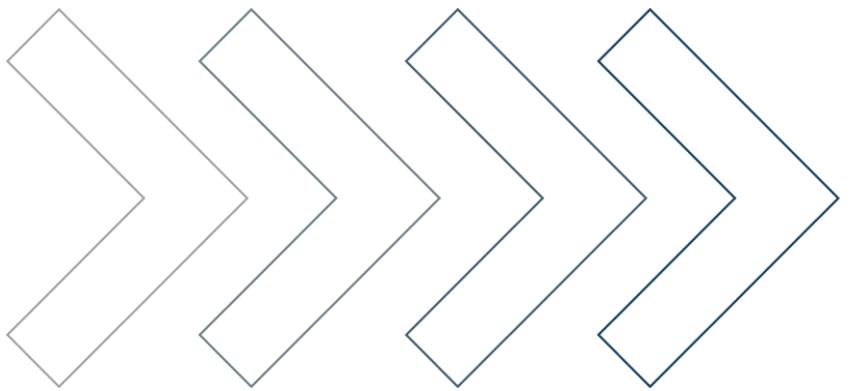


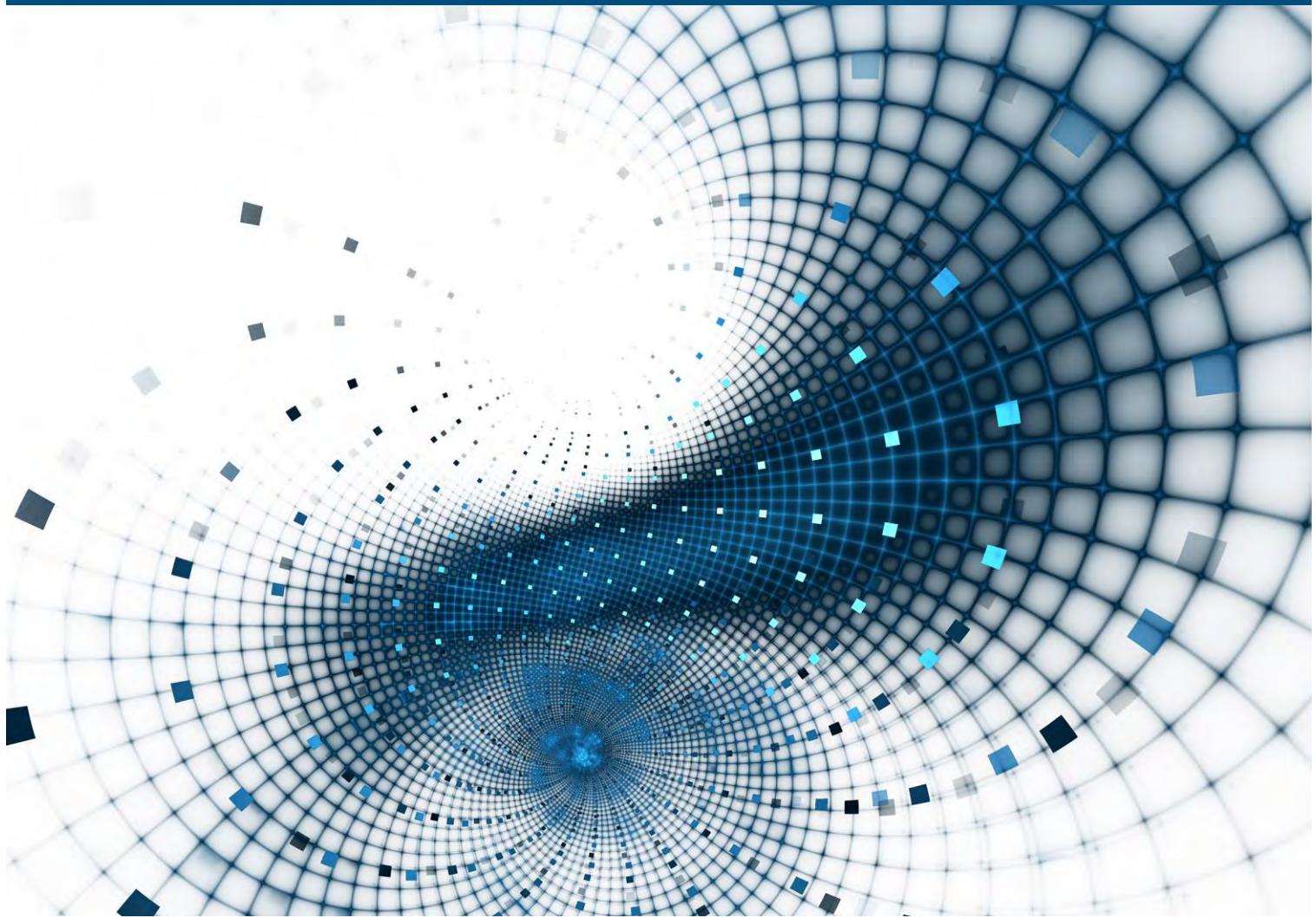


SOCIETY OF
ACTUARIES



Economic Scenario Generators

A Practical Guide



July 2016



Economic Scenario Generators

A Practical Guide

SPONSORS

Committee on Finance Research
Research Expanding Boundaries Pool

AUTHORS

Hal Pedersen, ASA, Ph.D.
Mary Pat Campbell, FSA, MAAA
Stephan L. Christiansen, FCAS, MAAA
Samuel H. Cox, Ph.D., FSA, CERA
Daniel Finn, FCAS, ASA
Ken Griffin, CFA, ASA, MAAA
Nigel Hooker, Ph.D., FIA
Matthew Lightwood, Ph.D., BSC (HONS)
Stephen M. Sonlin, CFA
Chris Suchar, FCAS, MAAA



Caveat and Disclaimer

The opinions expressed and conclusions reached by the authors are their own and do not represent any official position or opinion of the Society of Actuaries or its members. The Society of Actuaries makes no representation or warranty to the accuracy of the information.

Copyright © 2016 All rights reserved by the Society of Actuaries

CONTENTS

Introduction.....	6
Acknowledgments	6
Executive Summary	7
1. WHAT IS AN ECONOMIC SCENARIO GENERATOR?	7
2. WHY HAVE ECONOMIC SCENARIO GENERATORS EMERGED AS AN IMPORTANT TOOL TO SOLVE RISK MANAGEMENT PROBLEMS?	7
3. HOW DO ESGS HELP SATISFY REGULATORY REQUIREMENTS?	8
4. WHAT ARE THE SPECIFIC BUSINESS APPLICATIONS OF ESGS IN THE INSURANCE AND PENSIONS INDUSTRIES?.....	9
5. WHAT ARE THE ESSENTIAL FEATURES OF A COMPREHENSIVE ESG?	9
6. CONSIDERATIONS OF MODEL SPECIFICATION AND STYLIZED FACTS	10
7. MODEL CALIBRATION AND PARAMETERIZATION	11
8. VALIDATION.....	11
9. ARBITRAGE-FREE MODELING CONSIDERATIONS	12
10. RISK-NEUTRAL SCENARIOS	13
11. DEFAULT-FREE INTEREST RATE MODELS.....	14
12. CORPORATE BOND (YIELD) MODELS	15
13. EQUITY INDEX MODELS	15
14. INTERNATIONAL CONSIDERATIONS.....	16
Chapter 1: What Is an Economic Scenario Generator?.....	18
1.1 OVERVIEW OF ESG MODELS	18
1.2 ESG MODEL COMPONENTS.....	20
1.3 RELATIONSHIPS BETWEEN MODEL COMPONENTS	22
1.4 STRUCTURAL ESG MODEL COMPONENTS.....	22
1.5 CHAPTER SUMMARY.....	22
1.6 CHAPTER REFERENCES	23
Chapter 2: The Need for Economic Scenario Generators	24
2.1 SIMULATION VERSUS ANALYTICAL SOLUTIONS	24
2.2 ANALYTICS AND SIMULATION: AN ILLUSTRATION	26
2.3 HISTORICAL CONTEXT: THE BIG BANG AND EMERGENCE OF ESGS.....	29
2.4 ESG MODEL APPLICATIONS	31
2.5 HOW ESGS DIFFER FROM OTHER ECONOMIC MODELS.....	31
2.6 LIMITATIONS OF ESG APPROACHES.....	32
2.7 CHAPTER SUMMARY.....	34
2.8 CHAPTER REFERENCES	34
Chapter 3: The Role of ESGs in Satisfying Regulatory Requirements.....	35
3.1 APPLICATIONS OF ESGS IN RESPONDING TO EVOLVING REGULATORY REQUIREMENTS	35
3.2 SPECIFIC REGULATORY REQUIREMENTS WITH RESPECT TO ESGS.....	40
3.3 GENERAL REGULATORY ESG CONSIDERATIONS.....	41
3.4 CHAPTER SUMMARY.....	43
3.5 CHAPTER REFERENCES	44
Chapter 4: Business Applications of ESGs in the Insurance and Pension Industries	45
4.1 LIFE INSURANCE APPLICATIONS.....	45
4.2 PENSION AND RETIREMENT APPLICATIONS.....	49
4.3 PROPERTY AND CASUALTY (P&C) APPLICATIONS	51
4.4 CHAPTER SUMMARY.....	54

Chapter 5: What Are the Essential Features of a Good ESG?.....	55
5.1 REAL WORLD VERSUS MARKET CONSISTENT	55
5.2 ESSENTIAL FEATURES OF A GOOD ESG.....	56
5.3 ADVANCED ESG ATTRIBUTES	61
5.4 HOW GOOD IS AN ESG, REALLY?	62
5.5 CHAPTER SUMMARY.....	62
5.6 CHAPTER REFERENCES	62
Chapter 6: Considerations of Model Specification and Stylized Facts.....	63
6.1 STYLIZED FACTS.....	63
6.2 CONSIDERATIONS IN THE DETERMINATION OF STYLIZED FACTS	64
6.3 OVERVIEW OF COMMONLY MODELED ECONOMIC AND FINANCIAL VARIABLES.....	68
6.4 OVERVIEW OF THE EQUITY MARKETS	77
6.5 CHAPTER SUMMARY.....	77
6.6 APPENDIX: DATA SOURCES	78
6.7 CHAPTER REFERENCES	79
Chapter 7: Model Calibration and Parameterization	80
7.1 THE CALIBRATION PROCESS	80
7.2 CALIBRATIONS OF MARKET-CONSISTENT ESG MODELS.....	81
7.3 PARAMETERIZATION (CALIBRATION) OF REAL-WORLD ESG MODELS.....	83
7.4 METHODS OF PARAMETERIZING REAL-WORLD ESGS	85
7.5 CHAPTER SUMMARY.....	86
7.6 CHAPTER REFERENCES	86
Chapter 8: Model Validation	88
8.1 GENERAL CONSIDERATIONS IN DEVELOPING A VALIDATION SYSTEM	88
8.2 SPECIFIC CONSIDERATIONS IN VALIDATION OF REAL-WORLD ESGS	90
8.3 VALIDATION OF MARKET-CONSISTENT ESGS.....	93
8.4 VALIDATION: FINAL THOUGHTS.....	95
8.5 CHAPTER SUMMARY.....	96
8.6 CHAPTER REFERENCES	96
Chapter 9: Arbitrage-Free Modeling Considerations.....	97
9.1 ARBITRAGE-FREE FINANCIAL MODELS	97
9.2 APPLYING A PRACTICAL APPLICATION.....	99
9.3 IS THE ARBITRAGE-FREE FRAMEWORK NECESSARY?	100
9.4 MARTINGALE TESTS	101
9.5 CHAPTER SUMMARY.....	102
9.6 CHAPTER REFERENCES	102
Chapter 10: The Role of Risk-Neutral Scenarios.....	104
10.1 WHAT DOES “RISK-NEUTRAL” MEAN?	105
10.2 APPLICATIONS OF MARTINGALE TESTS IN PRACTICE	107
10.3 WHAT IS THE PURPOSE OF RISK-NEUTRAL SCENARIOS?.....	109
10.4 DECIDING BETWEEN REAL-WORLD AND RISK-NEUTRAL SCENARIOS.....	110
10.5 WHEN ARE REAL-WORLD AND RISK-NEUTRAL SCENARIOS USED TOGETHER?	111
10.6 CHAPTER SUMMARY	111
10.7 CHAPTER REFERENCES	112
Chapter 11: Default-Free Interest Rate Models	113
11.1 IMPORTANT PROPERTIES OF ROBUST INTEREST RATE MODELS	113
11.2 BINOMIAL INTEREST RATE MODELS	115
11.3 CONTINUOUS-TIME SINGLE-FACTOR MODELS OF THE TERM STRUCTURE.....	121
11.4 MORE ADVANCED MODELS FOR THE TERM STRUCTURE	134
11.5 CHAPTER REFERENCES	142
11.6 APPENDIX: MATHEMATICAL DETAILS FOR A GENERAL SINGLE-FACTOR MODEL.....	143

Chapter 12: Corporate Bond Models	151
12.1 DEFINITION AND KEY CONSIDERATIONS OF THE CORPORATE BOND MARKET.....	151
12.2 THE MECHANISMS OF CORPORATE BOND MODELING.....	153
12.3 STRUCTURAL/MERTON MODELS.....	154
12.4 REDUCED-FORM MODELS.....	155
12.5 RATINGS-BASED MODELS.....	157
12.6 OTHER ISSUES.....	161
12.7 CHAPTER REFERENCES	161
Chapter 13: Equity Index Models	163
13.1 GENERAL CONSIDERATION OF EQUITY MODELS.....	163
13.2 THE STANDARD BLACK-SCHOLES MODEL.....	165
13.3 THE HESTON MODEL.....	166
13.4 THE STOCHASTIC VOLATILITY WITH JUMPS (SVJ) MODEL.....	166
13.5 REGIME-SWITCHING MODEL	167
13.6 ECONOMETRIC MODELS	169
13.7 CHAPTER REFERENCES	169
Chapter 14: International Considerations.....	170
14.1 SINGLE-ECONOMY VERSUS MULTI-ECONOMY MODELING.....	171
14.2 DETAILS OF SIMILAR ASSET CLASSES CAN VARY SIGNIFICANTLY ACROSS ECONOMIES.....	172
14.3 DATA ISSUES IN MODELING GLOBAL ECONOMIES.....	173
14.4 THE IMPORTANCE OF CO-MOVEMENTS BETWEEN ECONOMIES.....	175
14.5 FOREIGN EXCHANGE (FX).....	179
14.6 CONSTRUCTING GLOBAL ECONOMIES FROM SINGLE-ECONOMY MODELS.....	183
14.7 CHAPTER REFERENCES	183
Annotated Bibliography	185
INTEREST RATES.....	185
CORPORATE BONDS AND CREDIT RATINGS.....	191
STOCK PRICE MODELS	192
INVESTMENT MODELS, ECONOMIC SCENARIO GENERATORS.....	194
Glossary	198
About the Society of Actuaries.....	Error! Bookmark not defined.

Introduction

An economic scenario generator (ESG) is a software tool that simulates future paths of economies and financial markets, and illuminates the nature of risk elements within the economy that drive financial variability. The analysis of a stochastic distribution of possible economic futures, which includes unexpected but plausible outcomes, is critical for testing a business model under a wide variety of economic conditions.

Applications of an ESG can range from limited tasks to broad analyses. For example, a small task might be the simulation of interest rates in order to understand the impact of interest rate changes on an investment portfolio. A broader role for an ESG could be supporting an enterprise risk management approach in order to identify and manage internal and external risks that might affect a company's financial stability and success. Some elements of regulatory regimes require firms to take into account the assumptions of internal business models in making business decisions. ESGs can provide greater transparency and specificity in satisfying such regulatory requirements.

An ESG can operate in the context of a "real-world" environment that captures market dynamics and risks in a way that a company will experience them. Or the context can be a "risk-neutral" environment that enables the pricing of cash flows that depend upon stochastic financial variability, such as would be found with investment guarantees in variable annuities.

This publication is intended to serve as an in-depth primer on economic scenario generators. The first half of the publication provides more general information on the nature of economic scenario generators—what they are, how they evolved and how they address regulatory and business needs in the insurance and pension industries. It also addresses essential features of a good economic scenario generator and provides some detailed guidance on important aspects of financial market model specification, model calibration and model validation considerations to assure that the ESG will produce simulation results that are relevant and sufficiently robust and that realistically reflect market dynamics.

The second half of the publication provides technical discussion of some of the essential financial markets aspects of a comprehensive ESG. Description of arbitrage-free considerations, the role of risk-neutral scenarios, and default-free interest rate models are explored with some of the underlying mathematics necessary to understand these considerations. Final chapters provide some important considerations relating to corporate bond models and equity index models, and how these considerations may extend to international and global interactions.

Each chapter begins by summarizing the specific sections of the chapter and ends with key takeaway points and additional references for further exploration. The publication closes with an annotated bibliography and a glossary of key terms used throughout the publication.

Acknowledgments

The authors would like to thank the following members of the Project Oversight Group for their participation and contribution to the development of this publication:

Dennis Radliff (Chair), David Cantor, Stephanie Ching, Russell Gao and Robert R. Reitano.

Executive Summary

1. WHAT IS AN ECONOMIC SCENARIO GENERATOR?

An economic scenario generator (ESG) is a computer-based model of an economic environment that is used to produce simulations of the joint behavior of financial market values and economic variables. Two common applications are driving the increased utilization of ESGs:

1. **Market-consistent (risk-neutral) valuation work for pricing complex financial derivatives and insurance contracts with embedded options.** These applications are mostly concerned with mathematical relationships within and among financial instruments and less concerned with forward-looking expectations of economic variables.
2. **Risk management work for calculating business risk, regulatory capital and rating agency requirements.** These applications apply real-world models that are concerned with forward-looking potential paths of economic variables and their potential influence on capital and solvency.

At its foundation, an ESG is concerned with simulating future interest rate paths, including yield curves. Financial markets operate within the context of the growth and volatility of economic markets. An ESG typically builds off core default-free interest rate modeling, then considers implications of corporate bond yields and returns that include default, transition behavior and stochastic spreads. Finally, other components, including equity markets, foreign-exchange considerations and economic components, such as inflation and GDP, may be considered.

An ESG model creates correlations through direct relationships with other simulated variables. Often, these variables and their interrelationships are modeled through a cascade structure to maintain model integrity. A cascade structure is a framework whereby each subsequent variable depends only on prior values of the variable and the values of variables that lie above them in the cascade structure.

2. WHY HAVE ECONOMIC SCENARIO GENERATORS EMERGED AS AN IMPORTANT TOOL TO SOLVE RISK MANAGEMENT PROBLEMS?

Historically, analytical solutions have been the standard methodology for the evaluation of risk metrics. An analytical process will result in a mathematically consistent and extendable solution, and statistical calculations can be performed quickly and accurately. However, analytical methods are most readily applicable to problems associated with variables where known or estimated underlying distributions exist. Much information about the economy is derived from empirical data. Analytical approaches are also less effective when dealing with complex actuarial analysis such as that required for asset liability management or financial risk management applications.

Simulation overcomes many of the restrictions that are imposed by the use of an analytical process and allows for the evaluation of complex dynamic financial systems such as those used for asset liability management and financial risk management.

With increasing sophistication of simulation modeling and advances in computing power, various economic and capital market models had evolved to be used for liability, investment and financial planning. The most familiar of these models was the mean-variance model. A mean-variance model has parameters that are identical to the assumptions that drive these models—namely, mean, standard deviation and correlations. This makes the calibration of these models relatively straightforward. Mean-variance models produced reasonable return projections, have nice mathematical properties and are easy to use and understand. However, mean-variance models used for sophisticated simulation modeling present problems when dealing with business applications that require analysis over multiperiod time horizons and when the analysis requires direct access to economic variables such as interest rates and inflation.

In the late 1980s, a new type of economic and capital market model utilized stochastic regression techniques within a cascade model structure to generate simulations of both economic variables (interest and inflation rates) and capital market returns (stocks and bonds). Over the years, the use of stochastic regression techniques has been overtaken by stochastic differential equation models, informed by the research within the field of option-pricing theory. Option-

pricing models proved to be more theoretically precise than regression-based models and allowed for both real-world projections and the pricing of contingent cash flows, both of which are critically important for a large subset of financial modeling applications. These new stochastic economic models were referred to by different names over the years, but the financial community eventually came to know them as “economic scenario generators.”

ESGs, while similar in form to other types of economic generation models such as pricing, forecasting and econometric models, differ from these models in that their primary goal is to generate a set of future economic scenarios that spans the range of possible future outcomes for use in risk management applications. Although economic scenario generators are extremely useful in gaining insight into future financial risk and rewards, like any model, they have limitations. Modeling the future dynamics of the economy and financial markets present many challenges, such as accounting for extreme events and regime changes. It is imperative that users of these models understand the strengths and weaknesses of any particular ESG so they can make sure that the ESG being used is appropriate for the analysis being performed.

3. HOW DO ESGS HELP SATISFY REGULATORY REQUIREMENTS?

In most frameworks, the use of ESGs is included or considered in possible responses to evolving regulatory rules and principles, but the development of strict definitions or requirements for ESG applications is still emerging. Regulatory applications are used to test a company’s solvency and liquidity position (value of the available surplus assets at the valuation date) as well as in support of individual product line pricing and reserving (required surplus assets and liability reserves) by quantifying the variations due to the realization of a company’s risk exposures. The first use establishes a time zero valuation, which can be compared with a benchmark for required surplus assets to absorb future realization of the entity’s risk exposures.

The most advanced regulatory frameworks where ESGs may be included or considered are seen in the emerging principles of ComFrame (a common framework for international insurance supervision established by the International Association of Insurance Supervisors) and in developing rules in EIOPA (the European body representing the regulatory authorities of the European Union and its developing solvency rules embodied under Solvency II). In the references to risk management frameworks associated with these regulatory regimes, allowance is made for stochastic analysis that achieves a sufficiently wide range of outcomes, reflects the scale and complexity of risks, and provides the ability to determine a statistical distribution of results. ESGs are shown throughout this publication to be more effective in achieving these goals, but for now, the guidelines permit but do not require this level of analysis. In emerging Solvency II regulations and prescriptions, standard and internal models used for quantifying risk and establishing minimum capital requirements may make use of ESGs.

In the United States under the guidelines of the National Association of Insurance Commissioners, ESGs are only beginning to emerge as an interesting, potentially preferred solution to solvency and product issues. ESGs have made the most headway in life insurance reserving regulations at the product level, and they are beginning to be recognized as an alternative in emerging ORSA (Own Risk Solvency Assessment) requirements. The NAIC has its own ESGs that are primarily applicable to regulatory considerations for interest and variable market-sensitive products in life and annuities (see the American Academy of Actuaries generator hosted at the Society of Actuaries website, <https://www.soa.org/research/software-tools/research-scenario.aspx>). These ESGs develop only a few economic factors: U.S. government yield curves and a handful of equity indices.

Considerations relating to the application of ESGs for regulatory purposes may be considered across all jurisdictions. These include time horizon and risk standards, governance, calibration, consistency across unrelated entities, conflicts with internal models, principles-based capital calculations, and expertise. ESG output used for regulatory purposes may not be appropriate to use for managing the business, as ESG model type and calibration may differ widely due to differing goals between regulators and insurers.

4. WHAT ARE THE SPECIFIC BUSINESS APPLICATIONS OF ESGS IN THE INSURANCE AND PENSIONS INDUSTRIES?

ESG model requirements differ for different business uses.

Applications of ESGs for life insurance liabilities are primarily focused on the interaction of interest rate changes and policyholder behavior regarding lapses and other optionality. Life insurance enterprise and product results are determined by the interaction of investment performance on assets built up from the collection of premiums, and the payout of liabilities based on events of mortality, morbidity (in health or disability) or policy surrender or annuity payout. Because of the complexity of the interaction of these factors over an extended time horizon, an ESG provides the type of comprehensive tool that is necessary to understand both the range of potential outcomes and the likelihood of scenarios. Representative applications in life insurance best understood through the use of an ESG include life liability valuation, effective duration analysis, stress testing, economic capital (EC) and strategic asset allocation (SAA).

In addition to similar applications discussed in the life insurance section, several specific applications in the pension industry are best understood through the use of an ESG. These applications include liability-driven investments, pension-funding sensitivities and pension risk transfer considerations. Liability-driven investment (LDI) is an investment strategy that defined benefit pension plans use to better match their assets with their liabilities and has become a popular technique over the past decade. Testing the effectiveness of this strategy and other pension investment strategies can be done with the use of an ESG. Pension plan sponsors are also using ESGs to project statutory funding requirements and to evaluate different funding strategies (e.g., the impact of a one-time large upfront contribution relative to more level or flat contributions). Real-world scenarios are needed to determine "likely" paths that the sponsor may have to navigate, and also to model the financial health of the sponsor. In risk transfer considerations, ESGs can be useful to sponsors looking to compare their options: choosing to freeze the plan and/or implementing LDI (whether plan is open or frozen) against options that involve transferring the risks to third parties.

Applications of ESGs in property and casualty insurance are more focused on the impact of inflation on liabilities and assets, and on economic cyclical characteristics affecting both exposures and policy pricing. Property and casualty products have different characteristics of liability development, with many casualty products involving significant time lags in settlement or even discovery. Property products are often volatile based on factors such as weather events, crime trends and economic supply-and-demand characteristics that may require greater liquidity reserves on the asset side. Because of the complexity of the interaction of these factors over an extended time horizon, an ESG provides a much more comprehensive tool for understanding the range of potential outcomes. Today an ESG is considered a necessary component of enterprise risk modeling for a P&C insurer.

5. WHAT ARE THE ESSENTIAL FEATURES OF A COMPREHENSIVE ESG?

It is generally expected that minimum requirements of an ESG would include the production of simulation results that reflect a relevant view of the economy and certain financial variables, the inclusion of some extreme but plausible results, and the generation of scenarios that embed realistic market dynamics. A comprehensive ESG has a sound foundation for the way the models are built and the way the variables are interrelated; it balances practicality and completeness.

The scope of an ESG application can vary widely. An ESG should be able to accommodate a broad range of applications involving both real-world and market-consistent scenarios, and use consistent models across both modes. Market-consistent scenarios are concerned with mathematical relationships within and among financial instruments. Real-world scenarios are concerned with forward-looking potential paths of economic variables and their potential influences. The simulation of interest rates and the portfolio investment strategy testing are examples of real-world simulation. The valuation of guarantees is an example of market-consistent simulation. The measurement of the effectiveness of a hedging program will often involve the application of both real-world and market-consistent simulations.

A comprehensive ESG strikes a balance between practicality and the ability to capture the critical features of the historical record. It also provides a suite of models that is sufficient to capture the asset classes and economic

variables that are of the greatest importance to the risk profile of a given firm, and provides additional modeling capabilities so that the available asset classes and economic variables can be expanded in a consistent manner to meet more specialized user requirements. It must meet the requirements of regulators and auditing firms and provide sufficient simulation detail for extensive validation. A comprehensive ESG is capable of accommodating many types of calibration views across a wide range of benchmarks while producing simulation results that reflect a view relevant to historical facts and that embed realistic market dynamics. It must produce extreme but plausible outcomes and also permit recalibration to different attributes to provide alternate sets of extreme but plausible events. Finally, a comprehensive ESG is computationally efficient and numerically stable.

6. CONSIDERATIONS OF MODEL SPECIFICATION AND STYLIZED FACTS

Stylized facts refer to generalized interpretations of empirical findings that provide a basis for consistent understanding of markets or economic drivers across a wide range of instruments, markets and time periods. Analysis of historical data is commonly used as the basis for determining stylized facts and setting calibration targets; however, stylized facts can also be based on expert judgement. Stylized facts are important in guiding the design of an ESG in that they help establish and prioritize the properties that the ESG model must have to be useful for a given application. The historical record of economic and financial markets is an indispensable guide to the dynamics that govern ESG model simulations. Detailed knowledge of these dynamics is essential for setting ESG model calibration targets and understanding strengths and weaknesses of various ESG model frameworks.

It is natural to summarize financial market variables in terms of their averages, standard deviations and correlations such as in a mean-variance framework. These summary statistics tell a good bit of the story, but they do not inform the subtle but important aspects of how markets are experienced through time. More advanced applications such as those used for pricing and risk management typically require additional specifications that may include information related to distributional shapes (fat tails), pathwise behavior (how variables move over time) and the ways characteristics of modeled variables change under different economic environments.

Stylized facts might include the following examples:

- Yields for longer-maturity bonds tend to be greater than yields for shorter-maturity bonds.
- Monthly fluctuations in short-maturity yields tend to be larger than monthly fluctuations in longer-maturity yields.
- When short-maturity yields are low, longer-maturity yields are normally higher than the shorter-maturity yields (a standard shaped yield curve). When short-maturity yields are high, longer-maturity yields are often lower than shorter-maturity yields (an inverted yield curve).
- Interest rates can be negative.
- Corporate credit spreads are wider for lower-credit-quality instruments. However, credit costs represent only a fraction of the spread on corporate bonds, which suggests that some portion of corporate bond spreads is due to factors other than credit costs (e.g., liquidity).
- There is a tendency for corporate credit spreads to fluctuate more during recessionary periods.
- The probabilities of default will fluctuate with general economic conditions and firm- or industry-specific conditions.
- Equity returns exhibit both higher expected returns and higher volatility than fixed-income returns.
- The volatility of equity returns fluctuates significantly over time.
- Correlations between modeled economic and financial market variables are not stable over time and can depend on whether monthly, quarterly or annual observations are being used.

If an ESG is part of an ongoing risk management program, the program usually requires that the data source be accurate, regularly updated and capable of automated download. The reason is that, in this use context, the ESG is likely to be recalibrated, validated and initialized to market conditions at regular intervals such as year-end or quarter-end. ESG practitioners are well served by having a wide range of data sources at their disposal. Historical data is the standard for setting stylized facts and targets that can be used for ESG model calibration and validation.

7. MODEL CALIBRATION AND PARAMETERIZATION

Users of ESG models need to incorporate a view of future market dynamics into their risk-modeling environment. The process of reflecting these views into an ESG is referred to as model calibration. More specifically, calibration is the process of setting the parameters of the equations within an ESG model to produce the distributions and dynamics (e.g., volatility, correlations, tail characteristics) of economic and financial variables required by the application for which they are being used.

As the model equations increase in complexity to account for more robust dynamics, the number of model parameters increases, and the linkage between the input parameters and the desired output statistics becomes less intuitive. The calibration process thus becomes increasingly challenging as additional variables are incorporated into the model. To make the calibration process practical, many ESG models use what is referred to as a cascade structure.

Different model calibrations may be required for different applications. Market-consistent valuation applications require ESGs to be capable of generating scenarios that can reproduce the observable prices of traded derivative instruments. ESGs used for these purposes need to adhere to strict mathematical properties that are designed to satisfy risk-neutral and arbitrage-free conditions. Because the model calibration process is designed to reproduce the prices of traded derivatives, the ultimate calibration depends on both the pricing date and the set of traded derivatives used to calibrate the model.

Calibration (also referred to as parameterization) of real-world ESG models requires users to make choices about the future economic environment they want to reflect in their risk analysis. Most risk management applications, for example, require ESGs to be capable of producing dynamics (e.g., volatility, correlations) that are representative of the possible future paths of economic variables. Because real-world parameterizations are forward looking, they require explicit views as to how the economy will develop in the future and thus require a significant amount of expert judgment to determine the veracity of the scenarios that result from the parameterization process. Some of the key decision points when parameterizing a real-world model include (a) selecting the appropriate steady-state levels, (b) determining the appropriate values for the initial conditions, (c) identifying the key parameterization targets or “stylized facts” that are necessary for the application, (d) controlling the expected reversionary paths of economic variables, and (e) identifying general assumptions.

Real-world model parameterization is an iterative process. Given the sheer number of parameters involved, a simple brute-force search is often impractical. Even many of the more traditional actuarial techniques, such as simple regression, are not effective, because the historical data for many economic variables violates one of the prime assumptions of these techniques: independent observations.

As a result, most parameterizations turn to more robust tools in this step. The first such tool is maximum-likelihood estimation (MLE), which involves calculating a likelihood function based on a possible set of model parameters and the observed data. Once the resulting probability function is determined, the set of parameters that maximize the likelihood function can be determined. The second key tool that can assist with parameterization is Kalman filtering, which involves consecutive cycles of predicting the state of an observed variable based on a model, comparing that prediction with the realized outcome in the historical observed data, and updating the parameters to achieve optimal predictive power.

8. VALIDATION

Validation ensures that the estimation of an ESG's parameters results in simulated behavior that is a good representation of the variable or market under consideration. Effective validation of an ESG requires comparing simulated output data with some predefined benchmark of acceptance criteria.

For a typical insurance or pension undertaking, the list of financial and economic variables that may be of interest is typically quite large. For this reason, the validation system and validation environment require careful design at inception, in order to organize the various data elements in an ordered fashion.

An automated validation system is preferable to manual validation. Validation should be repeatable and consistent through time. Before any data are analyzed or validation performed, it is helpful to form the acceptance criteria upon which the model output will be judged. This type of approach to validation, whereby the particular desirable features

of an ESG are based on analysis of a firm's risk exposures, is preferable to what might be called a "problem discovery" approach. In a problem discovery approach, a user first runs the ESG, creating a large output data set, and then tries to discover problems with the output.

Usually validation entails comparing the output of the ESG first with market data and finally with the acceptance criteria, which may be based on market data or a combination of market data and expert judgment. Considerations regarding the data to be used include the accuracy or degree of confidence that can be placed in the data, completeness (whether the data set is comprehensive), bias (data should be appropriately unbiased) and relevance. Effective validation will almost invariably need a human to visually review and make informed observations about the results. Using a validator with experience of the structure, limitations and usage of the model, who is able to combine both quantitative and qualitative judgments, is likely to be far more effective than relying on a particular set of rigid criteria or goodness of fit statistics to determine the robustness of the model.

Validation of real-world ESGs against market data falls into three distinct categories: point-in-time, ex-post or in-sample validation, and ex-ante or out-of-sample validation. Point-in-time validation involves checking that the time zero characteristics or initial conditions of the model adequately match those of the market. In-sample validation involves taking historical data and determining the extent to which the stylized facts of the data are captured by the ESG. Out-of-sample validation or back-testing involves looking at the output of a model from a point in time T and seeing how it performs at future time periods $T + t$. Out-of-sample validation helps to monitor the ongoing appropriateness of the model in light of what actually happens.

Validation should take account of a broad set of simulated variables with particular focus on the areas of greatest materiality. Validators should not restrict themselves to validating only the primary simulation variable (e.g., yield), but also derived quantities (e.g., returns and excess returns) in order to get a fuller picture of the quality of the ESG. Correlation and interdependency between variables is another important area to validate.

Unlike the validation of real-world scenarios, where there are many quantitative and qualitative possibilities for determining the quality of a scenario set, the validation techniques for market-consistent scenarios are fairly well defined: The model must only be validated for market consistency and risk neutrality. Validating market consistency involves first computing the prices implied by the model parameters and comparing the value of these prices to the current market prices. Validation and acceptance should concentrate on the areas of the fit where the highest risk exposure exists. The second important task is validating that the scenarios produced are in fact risk neutral. This means the expected return on a particular asset class is the same as the expected return on the risk-free asset. A standard test known as a martingale or $1 = 1$ test (i.e., a stochastic process in which the conditional expectation of the next value, given the current and preceding values, is the current value) is used to assess the extent to which this is true.

9. ARBITRAGE-FREE MODELING CONSIDERATIONS

Arbitrage is the ability to buy and sell securities in such a manner that the investor is able to generate a riskless profit. While arbitrage opportunities occasionally present themselves within the financial markets, they tend to be temporary discontinuities that are quickly traded away. Within ESG model applications, such as pricing and risk management, the existence of arbitrage is considered undesirable.

The importance of the theory of arbitrage-free financial models for ESG work is that it provides a mathematical framework that can be used to verify an ESG's consistency. The mechanism through which consistency is assured depends on the relationship between the cash flows of securities that are modeled and traded, and the prices of these securities.

An arbitrage is defined as an explicit trading strategy involving the simultaneous buying and selling of securities, possibly quite complicated, for which either the trader would receive a positive cash flow today with no further payment obligations during the remainder of the trade horizon, or the trader would enter into a trade at zero net cost and receive only nonnegative cash flows during the remainder of the trade horizon, with a positive cash flow occurring with probability greater than zero. The model is said to be arbitrage free if there do not exist any self-financing trading strategies that produce an arbitrage.

In an arbitrage-free model, securities producing identical cash flow streams are interchangeable from a pricing standpoint. Thus, derivative securities can be priced by first identifying a trading strategy that generates the same cash flows that are generated by the derivative security, and then using known prices of traded securities with the same cash flows to infer the price of the derivative security. The cash flow stream is hedged if it can be exactly replicated by dynamically trading in the securities market.

In an arbitrage-free model, prices of assets are discounted expectations across all future events, and arbitrage-free asset prices therefore embed all possible future behavior in today's prices.

Arbitrage-free models are a necessary requirement when applying an ESG to risk-neutral pricing problems. If the scenarios are not arbitrage free, then derivatives cannot be consistently priced. However, the dynamics that can be used to drive risk are constrained in an arbitrage-free model; those dynamics that can be applied may not lead to a reasonable pricing formula in today's experience.

Martingale tests are used to confirm that the scenarios generated by an ESG simulation are arbitrage free. The application of a martingale test is predicated on the assumption that risk-neutral output is available for the model to be tested. The simplest application of the martingale test is for assets that do not generate cash flows. If the scenarios are risk neutral, the time zero price of the asset should equal the average of the present value of the ratios of the risk-neutral prices and the values of the asset. When the risk-neutral output for a dividend-paying asset is to be tested, the martingale test can be based on the expected value of the discounted dividend stream over the horizon plus the discounted value of the asset at the horizon.

10. RISK-NEUTRAL SCENARIOS

Risk neutral means that an investor is indifferent to risk, or rather there is no explicit compensation for taking risk. Thus, in a risk-neutral world, an investor would be indifferent between two investment choices that have the same expected payoff but different risk profiles. In such a world, all investment alternatives, regardless of their risk profiles, would have the same expected return. It is typically agreed upon that in the real world, investors are not risk neutral but rather are risk averse. Thus, they require a higher expected return in order to be incented to incur additional risk to their invested capital. Nevertheless, the concept of risk neutrality is instrumental in the construction of a mathematically tractable environment that facilitates the pricing of liabilities, assets and derivative securities.

Understanding why the concept of risk neutrality is useful depends on the arbitrage-free pricing theory.

Risk-neutral theory provides a mathematical framework to find the price of an asset in an arbitrage-free context. The parameters of the stochastic process used to model the asset are calibrated so that the expected return from this asset over a short period of time is equal to the risk-free interest rate, and the expected discounted gains process for that variable is the same for all time horizons. In real-world risk management, calibration is done to obtain qualitative behavior that matches historical statistics or forward-looking views, whereas risk-neutral scenarios are not intended to conform to historical distributions but instead fit to today's market prices. That is, the dynamics of risk-neutral scenarios cannot be benchmarked against historical data.

The idea of risk neutrality is useful because it is intimately related to the pricing of assets and more generally cash flow streams, and provides a practical framework for computing prices for assets that are not generally traded in the market or subject to closed-form solution. It is also useful to be able to parameterize a model to exhibit qualitative behavior that is consistent with market norms.

Martingale tests are important to the application of risk-neutral theory in ESGs because technical results from martingale theory are needed to understand how to set the parameters to give risk-neutral behavior. Martingale tests arise in ESG practice because they are often required by regulators. The relationship between the price of a cash flow stream and the expectation of its discounted future cash flows for a risk-neutral scenario set provides a framework for the direct auditing of model output.

When an ESG is utilized for risk-neutral simulation, the focus of the calibration will be on the reproduction of derivative market prices. The idea is that the ESG should be calibrated to market price data as tightly as possible so that the pricing of insurance liabilities and other contingent claims that are not priced in the market will have prices that reflect liquid market prices as accurately as possible.

The general distinction is that risk-neutral scenarios are used for pricing and real-world scenarios are used to assess risk. But in real-world applications, risk-neutral scenarios will enter the picture when the real-world risk measurement exercise requires prices for liabilities or assets that do not have closed-form formulas. For example, in managing portfolios of mortgage-backed securities, the model needs to reflect complicated path-dependent cash flows that do not have a convenient closed formula to compute prices. Another example is the risk management and hedging of variable annuities that reflects policyholder behavior and the availability of many alternative investment products. An application such as this typically involves nested stochastic simulation, where the state of the economy is simulated under real-world measures to assess overall risk and to measure the effectiveness of a hedging strategy, and cash flows are priced using risk-neutral scenarios projected forward from each current node of the simulation.

11. DEFAULT-FREE INTEREST RATE MODELS

The default-free interest rate model is a key component of most ESG models. Its primary purpose is to generate the prices of risk-free bonds and for use in discounting liability cash flows. The collection of risk-free rates at various maturities makes up what is called the term structure of interest rates; this in turn allows for the construction of yield curves and the pricing of all default-free interest-rate-contingent cash flows.

Binomial interest rate models are based on a discrete time series and have been applied since the earliest days of interest rate modeling. Binomial models offer a computationally flexible approach for pricing a wide variety of cash flow structures. Multiperiod dynamics for the binomial model are described as a multiplicative binomial model.

Many processes can be used to generate default-free interest rate projections. Some of the most utilized models include the following:

1. **Continuous-time single-factor models of the term structure of interest rates.** These models are based on a stochastic differential equation process for the short rate (the instantaneous risk-free rate). This process incorporates an instantaneous drift of the short rate, an instantaneous standard deviation and a “noise” factor representing random shocks to the level of the short rate. The ratio of the expected return and the standard deviation leads to the market price of risk. This permits the pricing of discount bonds and valuation of interest-rate-contingent claims. Two formulations of short-rate processes that lead to discount bond prices are the Vasicek model and the Cox-Ingersoll-Ross model. Both involve mean-reversion and constant short-rate volatility, and in both, the market price of risk is assumed constant.
2. **Two-factor models of the term structure.** These models tend to provide more realistic models of interest rates. The two factors combine additively to produce the short rate. The market price of risk is constant but different for each of the two factors. The simulation of the term structure for the two-factor model then depends on the simulated two-dimensional trajectory of a two-dimensional process.
3. **Affine models.** These models assume that the dynamics of the term structure are driven by factors that are described by stochastic processes and represent the source of uncertainty of the models. Model parameters are commonly estimated using maximum-likelihood estimation. The affine class of models can be used under both risk-neutral and real-world measures because it has realistic behavior and accurate pricing. It allows well-defined estimation and calibration procedures, and the parameters can be efficiently estimated. Realistic correlations between yields within and across economies can be sustained. However, the models are linear, and in real data, nonlinearities can be important.
4. **Quadratic models.** These are classes of models where the short rate is a quadratic function of state variables. Quadratic models are more flexible in modeling nonconstant volatility and negative correlation among factors and can capture nonlinearity in the data. However, they are more difficult to fit targets of yields and returns, and the estimation procedure is heavy and time-consuming.
5. **The Nelson-Siegel model.** This model has some desirable properties and may perform better in linking the term structure to macroeconomic factors. It is widely applied for forecasting interest rates. In its most general form, it is not arbitrage free, but it is possible to restrict it to an arbitrage-free class. The model incorporates a level, a slope and a curvature to the term structure in a stochastic process that generates a range of plausible yield curves. The dynamics of the stochastic process need to be understood in the

context of the application, including the ability to capture historical yield curve behavior and extreme but plausible behavior.

12. CORPORATE BOND (YIELD) MODELS

A corporate bond is a security with a fixed schedule of payments (including coupons and a redemption amount), where the receipt of the scheduled payments is contingent on the issuing name being willing and able to pay both the coupons and principal components of the bond as scheduled. Pricing of a corporate bond will contemplate not only changes in the general level of interest rates, but also changes in the belief (ratings) in the potential for default and potential for recovery from default (migration risk and credit spread dynamics). Corporate bond models address the additional complications from migration risk and credit spread dynamics that will result in bond price fluctuations.

Corporate bond model approaches can be described in four general classes: structural models (or Merton models), reduced-form models (which are linked to survival distribution theory), ratings-based models, and hybrid credit models (which may be focused on specific credit applications).

Structural models are based on a model for the value of the firm that issued the debt to be priced. Simplifying assumptions are needed to get the theory into a tractable form, but the intuitions the model provides are valuable. Merton's original paper, "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates, (*Journal of Finance*, 1974)" remains a very readable presentation and an important source from which to understand this model.

Reduced-form models are based on point process models (or Poisson process mechanisms) that are focused on time of default and intensity of default (analogous to time of death and force of mortality in life insurance). These models offer no explicit mechanism for rating changes or ratings migration, so the class of models is used for rating certain classes of municipal bonds or sovereign bonds where rating changes are infrequent. Considerations in reduced-form models include the risk-free term structure, credit risk as an intensity process, loss given default, arbitrage-free pricing with a dividend, zero-recovery bonds and the role of risk premium.

Ratings-based models allow the user to prescribe a transition matrix and to model the changes in the spread of a corporate bond in response to changes in the underlying ratings. However, the trade-off for this feature is that idiosyncratic bond price behavior cannot be introduced, since any two bonds with the same rating, maturity and coupon will evolve to have the same price. To obtain a closed-form formula for these models requires several simplifying assumptions: It must be possible to evaluate the structure and factor processes, including interaction with factors that are driving the treasury model.

13. EQUITY INDEX MODELS

Equity index models are different from bond models. The value of a bond is known on its maturity date, but the price of an equity index is random in perpetuity. Equity indexes also are more volatile than most bond prices, and equity prices can exhibit sudden large changes over short periods that can be interpreted as jumps in the equity index price. This is not typical for bonds.

Realistic equity models are available that feature jump behavior, stochastic volatility and accurate return behavior. However, more robust models may have the disadvantage that they interfere with the concept of market completeness. An incomplete market means that sources of risk cannot all be hedged away by trading in the universe of available assets, and as a result, it is impossible to determine a unique price for cash flows that may be contingent on the equity index by arbitrage-free pricing considerations alone. Thus, additional assumptions may be needed to price certain derivatives or to model hedging of a variable annuity.

An equity index model produces total returns for the equity index, consisting of the combined effect of price returns and dividends. Models typically specify the ex-dividend price as one model component and the dividend yield as the other. Modeling dividend yield as a percentage of price, instead of cash dollar dividend, has the advantage of conforming to arbitrage-free restrictions, but dividend yield models do not reflect real-world market dividend payout behavior as well as a cash dividend model would. It is common that more than one equity index may be needed—for example, large-cap, midcap and small-cap returns. This requires modeling appropriate co-movements among the models, consisting of correlations and possibly even simultaneous jump movements across the indices. In applications

such as modeling for variable annuities and complex guarantees, the model must be capable of generating risk-neutral scenarios for the pricing of contingent cash flows.

The inclusion of observable economic variables does not significantly improve the explanatory power of an equity model, so most ESGs do not attempt to directly relate equity returns to macroeconomic effects.

The Black-Scholes model has been the traditional workhorse model for equity returns and associated derivatives pricing. The model assumes that the stock index price has a lognormal distribution with a constant price volatility parameter and a constant dividend yield.

The Heston model generalizes the Black-Scholes model by introducing stochastic volatility process in place of the constant volatility parameter. This has the advantage of generating more realistic equity return behavior and permits more accurate pricing of derivatives.

The stochastic volatility with jumps (SVJ) model generalizes the Heston model by introducing a jump component to the price level. The jump process can be taken to be of the compound Poisson type and usually incorporates a feedback where the volatility process affects the likelihood of a jump occurring. The model can produce very accurate equity return distributions but is mathematically complex and difficult to apply.

The regime-switching model builds on the Black-Scholes lognormal model with a random change in parameter dynamics. This can be interpreted as ongoing unobservable shocks to the economy. The simulation of the stock price is based on simulation of a (two-state) Markov chain with monthly transition probabilities together with normal random draw. The stochastic dividend yield process requires some approximation adjustments to maintain a tractable structure. The model is capable of significantly fattening the tails of stock price returns and can increase downside risk with no upside improvement, but there is no asset available to hedge the risk in a regime switch.

14. INTERNATIONAL CONSIDERATIONS

International considerations can enter into the application of an ESG for many reasons, such as inclusion of businesses that are multinational and achievement of a better model of today's international investment environment. Broadly speaking, moving to a multi-economy ESG imposes cross-economy calibration and validation requirements that are not a consideration in a single-economy context. The need for foreign-exchange rates is another distinction between single-economy and multi-economy ESGs. Finally, a multi-economy framework differs from single-economy modeling in the need for global correlation matrices and foreign-exchange processes that serve to link the constituent economies together into a global model.

Financial variables that appear on the surface to be similar can behave very differently in different economies. Since the objective of an ESG is to capture market behavior, the financial variables that an ESG simulates for a given economy depends on the actual market structures associated with that economy. ESG variables that appear in several economies may have materially different payment or risk characteristics despite sharing similar names.

A multi-economy ESG will likely encounter data issues in modeling global economies and associated challenges in setting calibration targets. One of the primary challenges in building multi-economy ESGs is that the availability of historical data varies greatly across economies. Some economies have long time series of data across a wide range of asset classes. Other economies may have short data histories for a small number of asset classes. Economies that are in transition may have available data series that need to be carefully interpreted in light of structural changes occurring in those economies.

Another challenge in developing a multi-economy ESGs is in the understanding of the co-movements among financial variables across economies and the limitations of what can be inferred from history. The simultaneous modeling of multiple economies requires that the components of each economy react sensibly as a global entity. There is no universally accepted way to measure co-movements of financial variables, but statistical correlation is one standard approach. Some of this can be achieved through correlation parameters, but some of this is structural and should be a part of the design of the dynamics of the global model. Two aspects of this general concept are the need to have a realistic proxy for global risk and a distribution of global scenarios that are suitable for understanding the diversification effects of various business strategies.

Exchange rate movements are another important consideration in multi-economy models. Foreign-exchange rates are the variables that complete the linkage between the component economies in a multi-economy ESG. Exchange rates are needed to understand the historical risk-return relationships between global asset classes, and exchange rates are needed to link the simulation results so that insurance risks and investment outcomes can be aggregated across economies.

The construction of an arbitrage-free multi-economy ESG requires that all models for traded assets fit together according to the restrictions that the arbitrage-free condition imposes. It turns out that it is possible to construct individual economies and then assemble these economies into a global arbitrage-free model, provided certain restrictions are imposed on the dynamics across economies. If the stochastic process that changes ESG variables within each economy (innovation terms) is not shared across other economies and if the innovation terms in the foreign-exchange processes are not shared across economies, then a global risk-neutral measure cannot be constructed.

Chapter 1: What Is an Economic Scenario Generator?

An economic scenario generator (ESG) is a computer-based model of an economic environment that is used to produce simulations of the joint behavior of financial market values and economic variables. The design and content of an ESG model can vary based on the needs of the specific application, but a general-purpose ESG could involve joint simulation of future bond prices, stock prices, mortgage-backed security prices, inflation, gross domestic product and foreign-exchange rates. The joint simulation includes not only simulated series of future values of each component, but their interactions as well. While an ESG does not include the direct modeling of liabilities that financial institutions would typically require in the management of their business, it would provide the fundamental economic output (interest rates, inflation, etc.) that may be necessary to project out these liability simulations.

This chapter presents a brief introduction of many concepts that will be further explored.

Section 1.1 gives an overview of ESG models. At its foundation, an ESG is concerned with simulating future interest rate paths, including yield curves. The section will explore how general-purpose ESG models consider economic and financial market returns associated with interest rate paths. An ESG is important in applications including enterprise risk management, where understanding external risks to an organization is critical. Finally, an introduction to the applications of market-consistent models and real-world models is presented. Implications relating to calibration and parameterization of these different applications are briefly discussed.

Section 1.2 describes common components that might be found within an ESG model. Financial markets operate within the context of the growth and volatility of economic markets. An ESG typically builds off core default-free interest rate modeling and then considers implications of corporate bond yields and returns. Finally, other financial markets, including equity markets and foreign-exchange considerations, as well as other economic components such as inflation and GDP, may be considered.

Section 1.3 discusses relationships between component models. Model calibration is introduced as a process for recognizing the interaction among financial and economic variables that would reasonably represent the future characteristics of those markets. This should produce a range of total return correlations that are consistent with historical data.

Section 1.4 outlines ESG model structural considerations, including examples of cascade structures. A cascade structure is a framework whereby each subsequent variable depends only on prior values of the variable and the values of variables that lie above them in the cascade structure.

1.1 OVERVIEW OF ESG MODELS

The heart of the ESG is its ability to generate future interest rate paths that can be used both in the discounting of future liability cash flows and in the projection of bond returns. In practice, it is rare that an ESG would be limited to the simulation of a single interest rate variable. More likely, most ESG models would project full yield curves for multiple term structures that often would include sovereign yields and corporate yield curves associated with multiple credit qualities. The number of yield curves and the underlying models for generating these future term structures would vary based on the requirements of the application for the ESG.

General-purpose ESG models tend to contain a broad set of modeled economic variables, which could include economic variables such as inflation, GDP and unemployment. Many ESG models also include the ability to model financial market returns. Common examples would include bond returns, stock returns, real estate returns and alternative-asset returns. The most sophisticated ESG models will even allow for the projection of derivative instruments. The need to service global institutions requires that some ESG models be able to extend the projection of economic and financial market models across multiple economies. These multi-economy ESGs will thus include currency exchange rates as part of their simulation capabilities.

Financial applications that require the use of an ESG are commonplace within the insurance, defined benefit pension, and banking industries. An economic scenario generator is integral to an enterprise risk management (ERM) approach

and is a critical component of the suite of models that analyze the external risks to an organization. An ESG feeds into the broader ERM framework and informs risk-based decision making. Essentially, an ESG is the basis for projecting economic and capital market scenarios for use in financial risk management applications.

Two common applications are driving the increased utilization of ESGs:

1. **Market-consistent valuation work for pricing complex financial derivatives and insurance contracts with embedded options (“market-consistent” models).** This application is mostly concerned with mathematical relationships within and among financial instruments and less concerned with forward-looking expectations of economic variables.
2. **Real-world models for risk management work in calculating regulatory capital and rating agency requirements.** These applications are concerned with forward-looking potential paths of economic variables and their potential influence on capital and solvency.

Market-consistent valuation applications require ESGs to be capable of reproducing the observable prices of traded derivative instruments in order to determine comparable prices for derivative instruments and insurance contracts with embedded options that are not traded but that require market valuation. ESGs used for these purposes need to adhere to strict mathematical properties such as risk-neutral and arbitrage-free conditions.

Because the market-consistent model calibration process is designed to reproduce the prices of traded derivatives, the ultimate calibration is dependent on both the pricing date and the set of traded derivatives used to calibrate the model. The validation associated with the model calibration is based on how well the model reproduces the market values of the universe of traded derivatives used to calibrate the model.

Risk management applications, in contrast, require that ESGs be capable of producing dynamics that are representative of the possible future paths of economic variables. Commonly referred to as “real-world” calibrations, they enable managers to ask what-if questions as they try to gauge the likelihood of future events and their impact on its business.

Because real-world parameterizations are forward looking, they require explicit views as to how the economy will develop in the future. Therefore, they require a significant amount of expert judgment to determine the veracity of the scenarios that result from the parameterization process. In practice, real-world calibrations are often parameterized to be consistent with historical dynamics of economic variables, although the long-term steady-state levels (expected means) associated with these parameterizations can differ from long-term historical averages in favor of current consensus expectations or specific user views. An example would be applying expert judgment to reflect a prolonged low-interest-rate environment as a result of central bank activities during the post-2008 environment.

Parameterizations of real-world ESG models require users to think about the future economic environment they want to reflect in their risk analysis. Some of the key steps involved in parameterizing a real world model include the following:

- Selecting the appropriate steady-state levels
- Determining the appropriate values for the initial conditions
- Identifying key parameterization targets for the application (stylized facts)
- Controlling the expected mean reversion path

The valuation of financial options embedded in life insurance and annuities, for example, would require a risk-neutral environment. In contrast, a property and casualty insurer might use an ESG with a real-world calibration to assess economic and regulatory risks.

A good set of simulated economic scenarios requires both a good set of model equations and an appropriate calibration. Carefully constructed model equations won't produce useful scenarios if the model is not well calibrated to reflect realistic behavior. Conversely, careful calibration will not produce a useful range of outcomes if the underlying model equations are not sufficiently robust to be able to reflect rare but possible occurrences. Ultimately, the quality of the ESG will generally be assessed by examining the simulated economic scenarios and validating them against the primary calibration targets.

1.2 ESG MODEL COMPONENTS

An ESG provides simulations of principal financial market and economic variables that in aggregate represent a necessary and sufficient picture of the behavior of these markets that are necessary to provide insight into the specific risk management questions being asked. Financial markets operate within the context of the economy (or global economy in many cases). Elements of growth, scale and volatility of economic components are essential to understanding characteristics of economic and financial market dynamics.

This section provides a brief description of some of the important economic and financial market components that are typically found within a general-purpose ESG model. Many of these economic and financial components are described and developed with some technical detail in later chapters of this paper.

A good ESG provides a suite of models that are sufficient to capture the asset classes and economic variables of greatest importance to the risk profile of a given firm. Many risk management applications will require modeling the following variables:

- Sovereign interest rates and returns
- Corporate bond yields and returns
- Equity returns
- Foreign-exchange rates

Some applications may require additional variables such as these:

- Inflation
- GDP
- Unemployment
- Mortgage-backed securities
- Covered bonds
- Municipal bonds (important in U.S. models)
- Derivatives

1.2.1 SOVEREIGN INTEREST RATES AND RETURNS

Interest rate modeling is at the core of an ESG. This is true because of the importance that interest rates play in discounting liability cash flows and in determining returns on fixed-income investments. Early financial models such as Wilkie (1986) developed approaches based on regression methods to project financial markets, securities or economic drivers. The need to produce more insight into interest rate dynamics has led to the use of more academically rigorous models, which now defines the current state of interest rate modeling within economic scenario generators.

Modeling interest rates can be further refined to consider yield curve modeling and the modeling of the dynamics associated with the yield curve. Considerations incorporating inflation would be necessary to reflect real interest rates in addition to nominal rates. Depending on the application, bond and coupon considerations may be important to gain the necessary insight into the risk brought on by interest rate dynamics.

Bond returns are a particularly important ESG component in the context of insurance and other financial institutions. Bond prices are a function of changes in sovereign interest rates. Typically the bond component of an ESG will include

sovereign zero-coupon bonds, sovereign coupon bonds, corporate and municipal bonds with varying degree of default risk, and mortgage-backed securities (MBS).

Financial institutions invest heavily in bonds. The U.S. life insurance industry, for example, invests the majority of its admitted assets in bonds. The preponderance of these bond holdings consists of corporate and structured securities.

1.2.2 CORPORATE BOND YIELDS AND RETURNS

The market for corporate bonds exhibits some interesting and diverse features. Pricing of a corporate bond contemplates not only changes in the general level of interest rates, but also changes in the probability of the potential for default (ratings) as well as the potential for recovery from default (migration risk and credit spread dynamics.) Corporate bond models consequently need to address the additional complications from migration risk and credit spread dynamics that will result in greater volatility in bond price fluctuations.

Corporate bond model approaches can be described in four general classes: structural models (or Merton models), reduced-form models (which are linked to survival distribution theory), ratings-based models and hybrid credit models (which may be focused on specific credit applications). Credit modeling is, however, in its relative infancy, and new research continues to produce more complete models of credit yields and corporate bond returns.

1.2.3 EQUITY MARKETS

Although insurers invest heavily in bonds, equities are also important to model within an ESG, because insurers often include an equity allocation as part of their investment strategy. In some cases, equities are crucial. For example, a life insurer may offer variable annuities or universal life insurance, which credit policyholders with an equity portfolio return subject to a minimum guaranteed return. In such cases, an ESG provides scenarios that include equity returns for use in managing these guarantees.

Pension plan sponsors also would have significant interest in the equity markets, as equity exposure has traditionally been a major investment choice underlying the assets backing their long-term retirement liabilities.

The Black-Scholes model was one of the first stochastic equity models. This model is widely used in pricing applications. There are many more general equity models that are designed to more appropriately model long-term behavior. These advanced equity models are more often used to model equity markets within ESG modeling systems.

1.2.4 FOREIGN-EXCHANGE RATES

An ESG component for foreign-exchange rates is important for financial institutions with international operations. An insurance company doing business in more than one country will likely require an ESG with foreign-exchange components. A large pension plan may invest globally, even though all plan participants are in its home country and paid in its domestic currency. Foreign-exchange rates would also come into play when a company wants to manage its global pension plans in a holistic or coordinated manner.

1.2.5 OTHER ECONOMIC AND FINANCIAL MARKET COMPONENTS

While interest rates are an essential component of an ESG, other economic variables, such as Inflation, GDP and unemployment, also may have a direct role in determining the dynamics of liability cash flows. Therefore, these additional economic components are often available within a general-purpose ESG.

Other financial market components may also be important or useful, depending on the application, and would be prioritized within the limits of the practicality of producing a workable model. A specific application may require financial market components in addition to sovereign and corporate bond returns such as mortgage-backed security returns, covered bonds and derivatives. In the United States, municipal bonds may be an important component. These may in turn depend on other financial market components.

1.3 RELATIONSHIPS BETWEEN MODEL COMPONENTS

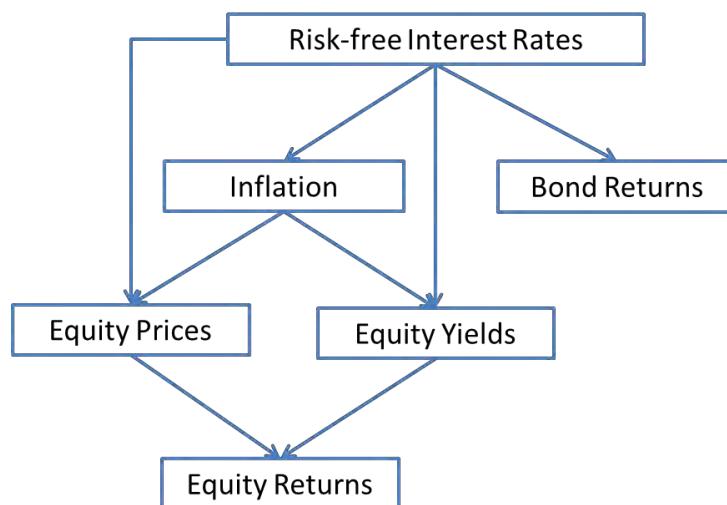
Each economic variable and financial market return modeled within an ESG needs to be projected through time and across multiple scenarios in a manner that, when the components are taken together, results in projections that reasonably represent the future characteristics and dynamics of the underlying variables being modeled. In addition they need to be projected in a manner that represents how these variables will move in relation to the movements of all of the other variables within the model. The process for meeting these conditions is called model calibration.

An ESG produces a range of total return correlations. Large changes in correlations can result from systemic effects such as the flight to quality during a financial crisis or the effects of Federal Reserve policies. An ESG embeds dynamics that affect correlations in order to adequately capture changes in investment opportunity sets. Correlations are important for capturing the risks associated with multi-asset class investing and for understanding the effects of diversification and concentration.

1.4 STRUCTURAL ESG MODEL COMPONENTS

An ESG model creates correlations through direct relationships with other simulated variables. The calibration process itself becomes increasingly challenging as additional variables are incorporated into the model. To make the calibration process practical, ESG models make use of what is referred to as a cascade structure. A cascade structure is a model framework design whereby each subsequent variable in the cascade structure depends only on prior values of the variable and the values of variables that lie above them on the cascade structure. This prevents changes to variables in the lower levels of the cascade from adversely affecting the dynamics associated with variables higher up in the cascade structure. Under a cascade structure, once a variable is appropriately calibrated, the calibration of subsequent variables lower on the cascade structure will have no impact on the previously calibrated variables. Figure 1.1 shows a cascade structure that might be used within an economic scenario generator.

Figure 1.1
CASCADE STRUCTURE OF A HYPOTHETICAL ECONOMIC SCENARIO GENERATOR



1.5 CHAPTER SUMMARY

1. An economic scenario generator is a computer-based model of an economic environment that is used to produce simulations of the joint behavior of financial market values and economic variables.
2. Interest rate modeling is at the core of an ESG because of the importance of interest rates in discounting liability cash flows and in determining returns on fixed-income investments.

3. Interest rates are closely associated with default-free bonds and may be further refined to consider dynamics of the yield curve and interaction with inflation.
4. Other financial components likely considered in a robust ESG may include gross domestic product, inflation, unemployment rates and foreign-exchange rates.
5. Key financial market and economic variables are modeled in simulated series, and importantly the interaction among the variables also is modeled. Modeling in an economic scenario generator can be performed on a market-consistent basis or real-world basis. Each has application to the insurance and pension worlds and to understanding financial markets.
6. Some key financial market variables particularly important to insurance and pensions are bonds (and related interest rates), including corporate and asset-backed bonds, and equities.

1.6 CHAPTER REFERENCES

Black, Fischer, and Myron Scholes. 1973. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy* 81 (3) (May–June): 637–54.

Wilkie, A. D. 1986. A Stochastic Investment Model for Actuarial Use. *Transactions of the Faculty of Actuaries* 39: 341–403.

Chapter 2: The Need for Economic Scenario Generators

Having described economic scenario generators (ESGs) and their likely components in the first chapter, we will now discuss why and how ESGs emerged, with particular emphasis on the insurance industry environment.

The increased emphasis on risk and capital management within the financial services industry has created the need for large-scale risk management models that consider risk exposures that affect both sides of the balance sheet. While risk modeling is not new, the use of ESGs to facilitate large-scale financial risk management simulation is a fairly recent phenomenon.

This chapter will discuss the role that ESGs play in financial risk management and why these systems have gained increasing importance.

Section 2.1 will describe the application of simulation modeling to address problems that do not lend themselves to analytical solutions. Analytical solutions best apply when the underlying distributions are known or can be estimated. Much information on the economy is derived from empirical data that often does not fit to a closed-form distribution. Other problems arise in complex joint distributions, discontinuities and mapping to other variables. With improvements in technology and computing power, simulation approaches emerged as a computational alternative.

Section 2.2 provides an illustration contrasting the use of analytical and simulation approaches in calculating insights into complex problems. In an insurance environment, there is often a need to measure risk exposure resulting from the combination of a number of variables, including frequency and severity of claims and the total amount of claims. The analytical effort to model the compound distribution of these variables is far from trivial. Simulation provides a more accurate estimation for such complex modeling.

Section 2.3 will examine the genesis of ESGs as an extension of simulation modeling. The genesis of what is now referred to as economic scenario generators can be traced to the confluence of three developments: advances in computing power; financial volatility following the collapse of the Bretton-Woods system of fixed international exchange rates and the subsequent 1973–74 stock market crash, which put pressure on the banking and insurance industry to come up with better financial markets modeling approaches; and the introduction of option-pricing models that proved to be more precise than regression-based models.

Section 2.4 will give some examples of how ESGs are being used to drive financial risk management applications. Important examples (among others) include asset liability management, investment strategy, regulatory compliance, hedging and reinsurance pricing.

Section 2.5 will distinguish ESGs from other types of economic models such as pricing, forecast and econometric models. The key difference is that economic scenario generators focus on economic risk rather than attempting to give insight into how the economy works or to predict the future level of an economic variable.

Section 2.6 provides a discussion of some of the challenges and limitations of using ESGs as part of a simulation framework. Key among these limitations are the issue of sample error, extended processing time, and the need for substantial amounts of data and advanced technical expertise.

2.1 SIMULATION VERSUS ANALYTICAL SOLUTIONS

Historically, analytical solutions have been the standard methodology for the evaluation of risk metrics. An analytical solution, from a mathematical perspective, is the process of finding a solution using a methodology involving algebra and calculus to produce an expression for the solution in a closed-form equation. There are many reasons why the use of an analytical solution is desirable. Foremost among them is that an analytical process will result in a mathematically tractable solution. Expressing a solution as a closed-form equation makes it possible to perform many statistical calculations quickly and accurately.

2.1.1 IMPEDIMENTS TO ANALYTICAL SOLUTIONS

While an analytical solution is preferable to a simulation approach, it often imposes substantial restrictions on the analysis being performed. Examples include situations with the following conditions:

- The underlying distribution of data is unknown or unknowable.
- The problem involves the use of complex joint distributions.
- Discontinuities must be incorporated into the analysis.
- Results require translation or mapping to the desired output.
- Multiperiod projections are required.

2.1.2 LIMITATIONS OF ANALYTICAL SOLUTIONS

Analytical methods are most readily applicable to problems associated with variables where known or estimated underlying distributions exist. However, much information about the economy is derived from empirical data. Serious difficulties are often encountered when attempting to fit this data to a known underlying distribution. Since the analytical process requires that the solution be ascertained from a problem that can be stated in closed form, empirical distributions cannot be used directly. Thus, analytical methods are not feasible when dealing with variables whose distributions cannot be expressed in closed form or can only be expressed approximately.

Analytical solutions are also less effective when dealing with complex actuarial analysis such as that required for asset liability management or financial risk management applications. This is true even when the variables used in these applications are assumed to follow well-defined distributions. This is due to the mathematical complexity that arises when working with joint distributions of stochastic dynamic processes, and aggregation, where the aggregated variables follow different distributional forms.

Discontinuities, a problem that often arises when dealing with economic and capital market variables, present yet another major stumbling block to those attempting to find an analytical solution. Discontinuities also arise within institutional liabilities as a result of retentions and limits associated with insurance policies and with data clustering. These types of discontinuities often make specifying and solving the problem in closed form difficult or impossible.

Finally, especially in the cases of asset liability management and financial risk management, the results of the analysis are often expressed in terms that are different from the terms used to describe the underlying variables. This requires the use of mapping functions that do not always follow equations that can be readily specified in closed form. For example, a particular investment might be stated in units of total return and assumed to follow a lognormal distribution. The resulting output, however, might need to be assessed in terms of financial statement line items, such as income or surplus. Thus, a transformation that takes into account a complex set of accounting and regulatory rules would need to be applied to the investment returns. This is not typically feasible using an analytical approach.

2.1.3 ANALYTICS SUPERSEDED BY APPROXIMATION TECHNIQUES

As financial modeling applications increased in complexity, pure closed-form solutions gave way to approximation techniques. Approximation techniques were heavily relied on before the invention of computers, as they balanced the competing interests of accuracy and computational convenience. The advent of the computer and the subsequent major advancements in computing accessibility, power and speed has made the use of simulation methodologies a much more practical and powerful alternative for the evaluation of problems involving institutional assets and liabilities.

2.1.4 SIMULATION EMERGES AS A COMPUTATIONAL ALTERNATIVE

Simulation is the representation of the operations or features of one process (or system) through the use of another. It is a standard process used in the field of engineering, where it reduces the time and cost associated with the production of real prototypes and experimental testing. Simulation conforms nicely to the needs of computational programming. Simulation also overcomes many of the restrictions imposed by the use of an analytical process and allows for the evaluation of complex dynamic financial systems such as those used for asset liability management and financial risk management.

Simulation requires the use of random samples to find a solution to a mathematical problem. This statistical sampling technique is often referred to as a Monte Carlo process. Since simulation involves random sampling from a set of uncertain possibilities, the simulation process requires (1) the specification of the set of uncertain possibilities, (2) the assigned probability estimates relating to the set of possibilities and (3) a random-number generator.

2.2 ANALYTICS AND SIMULATION: AN ILLUSTRATION

To show how quickly an analytical process can grow in complexity, we will examine an evaluation that might be done within an insurance institution.

Claim payments associated with natural catastrophes are often calculated as a joint distribution of frequency and severity, where frequency measures the number of claims or events (e.g., hurricanes and earthquakes) and severity measures the magnitude of loss of an individual claim or event.

As an example, consider the need to measure the potential risk exposure arising from a particular line of business. A company will be concerned not only with the expected or average amount of claims, but also what the total probable amount of what the claims might be.

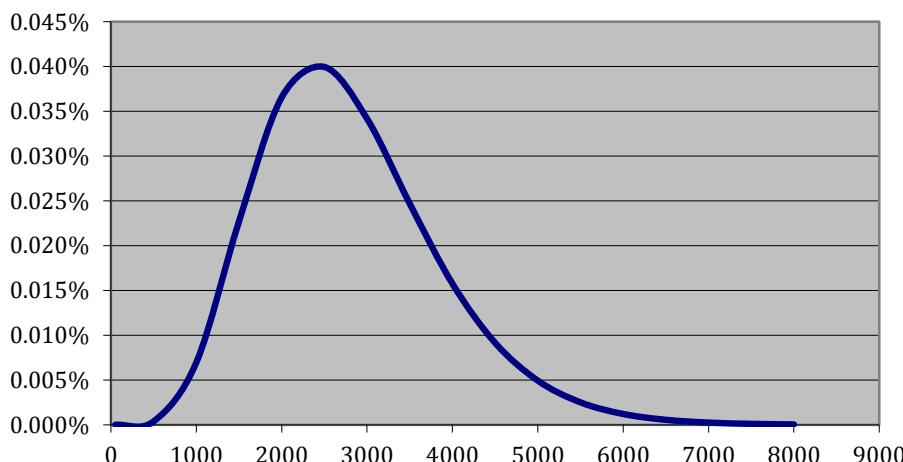
2.2.1 ANALYZING THE CHARACTERISTICS OF A SINGLE DISTRIBUTION

Assume that for a given event, the size of loss distribution (severity) for the line is best described by a gamma distribution. The gamma distribution (Γ as a scalar) with parameters shape = λ and scale = α has density

$$f(x) = \lambda^\alpha / \Gamma(\alpha) x^{(\alpha-1)} e^{-(\lambda x)}$$

Further, assume the parameters of the gamma distribution to be $\alpha = 7$ and $\lambda = 0.0025$. Figure 2.1 graphs this distribution.

Figure 2.1
GAMMA DISTRIBUTION FOR $\alpha = 7, \lambda = .0025$



Given this information, closed-form equations can be used to calculate some basic statistical measures, such as the mean and standard deviation:

$$\text{Mean} = \mu = \alpha/\lambda = 2,800.0$$

$$\text{Standard Deviation} = \sigma^2 = \alpha/\lambda^2 = 1,058.3$$

Further, using a Taylor series expansion makes it possible to calculate a numerical estimate for the percentile distributions that gives the results shown in Table 2.1.

Table 2.1
ANALYTICALLY DERIVED PERCENTILES FROM A GAMMA DISTRIBUTION WITH $\alpha = 7$, $\lambda = 0.0025$

Percentile	Exceedance Probability	Loss Severity (\$, 000s)
90th	1 in 10	4,213
95th	1 in 20	4,737
99th	1 in 100	5,828
99.6th	1 in 250	6,401
99.8th	1 in 500	6,818

Table 2.1 shows the loss severity associated with a single event at various percentile levels. The 90th percentile loss from a single event would generate \$4,213,000 in claims, assuming the severity of loss followed a gamma distribution with the stated assumptions. Thus, there would be a 1 in 10 chance of losses exceeding \$4,213,000 (90th percentile) and a 1 in 500 chance of losses exceeding \$6,818,000 (99.8th percentile).

Alternatively, simulation could have been used to come up with estimates for each of these same statistical measures of probability of losses. Using 10,000 Monte Carlo simulations generated the results for each of five separate trials (each comprising 10,000 simulations) shown in Table 2.2.

Table 2.2
ANALYTICAL VERSUS SIMULATED LOSS STATISTICS FROM A SINGLE OCCURRENCE

	Loss (\$, 000s)					
	Analytic Result	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5
Mean	2,800	2,784	2,810	2,797	2,793	2,798
Standard deviation	1,058	1,048	1,049	1,055	1,054	1,064
90th	4,213	4,174	4,200	4,208	4,195	4,246
95th	4,737	4,747	4,711	4,704	4,703	4,783
99th	5,828	5,785	5,803	5,833	5,855	5,874
99.6th	6,401	6,415	6,324	6,433	6,288	6,433
99.8th	6,818	6,722	6,757	6,841	6,735	6,782

Table 2.2 shows that the analytical solution for the mean loss from a single event is \$2,800,000, while the five simulation trials generated estimates ranging from a low of \$2,784,000 to a high of \$2,810,000. The analytical result at the 1 in 500 level suggests losses would not exceed \$6,818,000, while the simulation trials suggest at this level that losses would not exceed a range from \$6,722,000 to \$6,841,000. Note that the trials and simulations illustrated here

are all based on the same underlying distribution and parameters. The different trial estimates are the result of error or noise in the simulation process, reinforcing the fact that samples from a known distribution never exactly match the theoretical outcome.

Nevertheless, the differences in the simulation outcomes are relatively small, assuming that the parameters and distribution equation, as selected, accurately represent the underlying process. If this were the extent of the analysis, then the analytical approach would be preferable, due to considerations of tractability, speed and accuracy.

2.2.2 THE PROBLEM OF INCREASED COMPLEXITY FROM JOINT DISTRIBUTIONS

However, this analysis only deals with loss information related to a single event. What if calendar year loss information were required for use in budgeting and strategic planning? In addition to the event severity, the frequency of events would need to be considered.

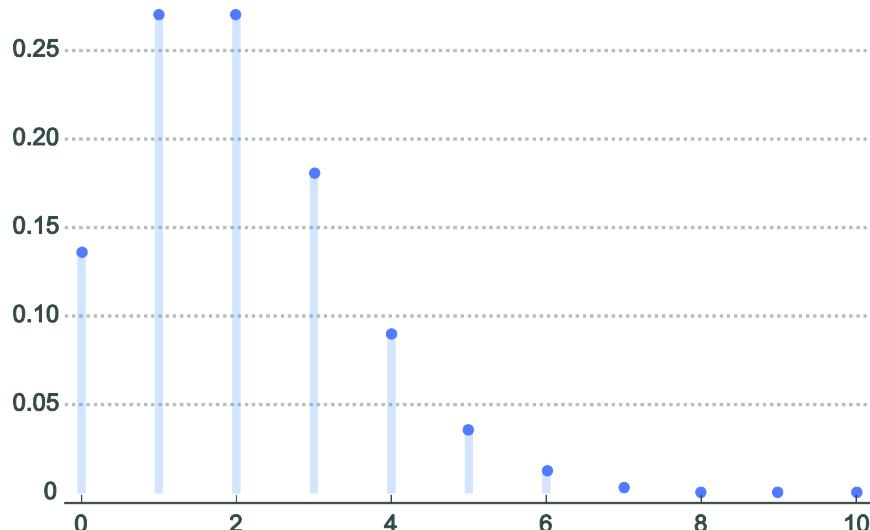
Assuming that each event is independent and identically distributed, it is easy to calculate the relevant statistics for any discrete number of events. However, the frequency of events is not typically certain and often is assumed to follow a separate stochastic process.

Continuing with our example, assume that a Poisson distribution best describes the frequency of events over a given calendar year. The Poisson distribution with parameter μ has density

$$f(x) = (\mu^x e^{-\mu}) / x!, \text{ where } x = 0, 1, 2, \dots$$

Here, x represents the number of events occurring over a given interval of time, and μ represents the expected number of such events occurring.

Figure 2.2
POISSON DISTRIBUTION FOR $\mu = 2$



If an average of two events are expected per annum (around a distribution that could include no events or a large number of events), we would need to evaluate the compound distribution made up of the product of a Poisson distribution (frequency) and a gamma distribution (severity), as described before. While the mean and standard deviation of this compound distribution can still be solved analytically, it is by no means trivial to solve this distribution for the various percentile loss levels.

The complexity of this compound distribution suggests the use of numerical approximations or simulation to get this information. Table 2.3 shows the results for each of five separate trials using a sample set of 10,000 simulations.

Table 2.3**ANALYTICAL VERSUS SIMULATED CALENDAR YEAR LOSS STATISTICS**

	Loss (\$, 000s)					
	Analytical Result	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5
Mean	5,600	5,610	5,649	5,656	5,569	5,542
Standard deviation	4,233	4,241	4,230	4,247	4,227	4,236
90th	*	11,425	11,287	11,334	11,331	11,315
95th	*	13,508	13,480	13,502	13,410	13,402
99th	*	17,906	17,896	17,996	17,798	17,817
99.6th	*	20,202	20,242	20,259	19,901	19,945
99.8th	*	21,925	21,644	22,226	21,781	21,635

*Actual closed-form analytical solution not practical for tail percentile calculations.

Table 2.3 shows that the analytical solution for the mean calendar year loss is \$5,600,000, while the five simulation trials generated estimates ranging from \$5,542,000 to \$5,656,000. But an actual closed-form analytical solution for the various percentile levels is not practical, whereas the simulation approach provides management with estimates that the 2 in 1,000 calendar year loss can range from \$21,635,000 to \$22,226,000. Thus, simulation is a superior approach for providing these percentile estimates.

2.2.3 FURTHER COMPLEXITY AS ADDITIONAL DISTRIBUTIONS ARE INTRODUCED

The impact of changes in economic variables such as uncertainty in inflation on a company's loss exposure also may be relevant. Since inflation can follow yet another, different, stochastic process, the analysis becomes even more complicated. Simulation techniques thus become the only feasible method for performing the necessary analysis.

2.3 HISTORICAL CONTEXT: THE BIG BANG AND EMERGENCE OF ESGS

Economic scenario generators go beyond analytical and simulation modeling, responding to additional problems of data and complexity as financial modeling requirements became more sophisticated. The genesis of what is now referred to as economic scenario generators can be traced to the confluence of three major and largely independent global developments:

1. Advances in computing power, which facilitated the increased use of simulation modeling
2. The collapse of the Bretton-Woods system and the subsequent 1973–74 stock market crash
3. The introduction of option-pricing theory

2.4.1 ADVANCES IN COMPUTING POWER FACILITATES THE USE OF SIMULATION MODELING

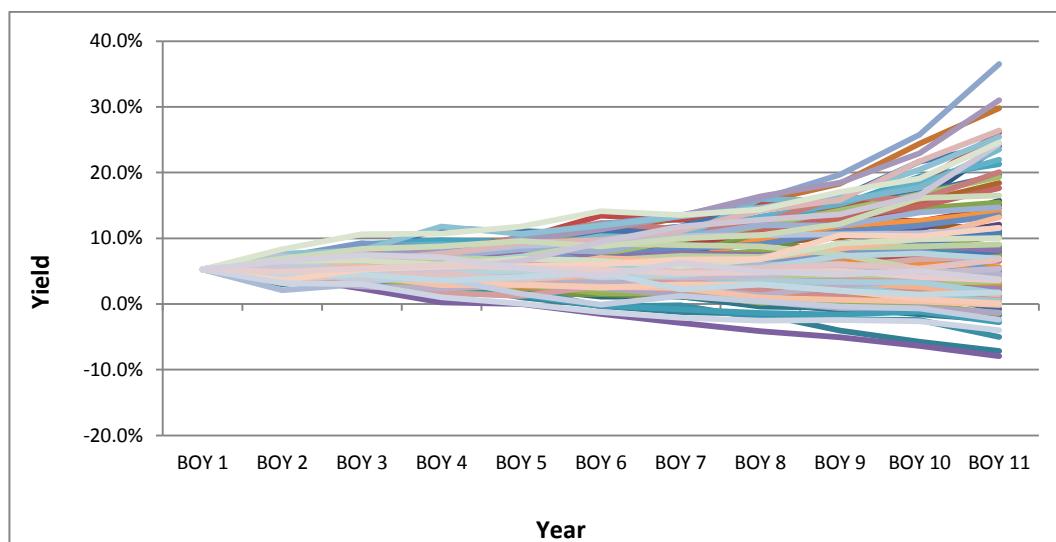
Through the 1970s, with increasing sophistication of simulation modeling and advances in computing power, various economic and capital market models had evolved to be used for liability, investment and financial planning. The most familiar of these models was the mean-variance model popularized by Harry Markowitz in his seminal work, *Portfolio Selection: Efficient Diversification of Investments*. Mean-variance models produced reasonable return projections, had nice mathematical properties and were easy to use and understand. However, mean-variance models, even used within sophisticated simulation modeling, presented problems when dealing with business applications that required

analysis over multiperiod time horizons and when the analysis required direct access to economic variables such as interest rates and inflation. The technical problem of handling multiperiod extensions was addressed within the mean-variance model by making the simplifying assumption that capital market returns were independent from period to period. However, the need to generate economic information in addition to capital market returns presented a much tougher challenge.

It is a relatively straightforward exercise to calculate Treasury bond returns using a mean-variance model and then to solve for the implied yields that would be required to generate the projected returns. When this is done, it can be shown that the underlying implied yields are nonstationary. That is, the minimum and maximum implied yields underlying the projected bond returns become progressively larger and smaller over time, ultimately moving toward positive and negative infinity (see Figure 2.3). This result clearly pointed out a major flaw within the mean-variance model.

Figure 2.3

IMPLIED 10-YEAR TREASURY YIELD



Prepared by Conning, Inc. Source: Conning's GEMS® Economic Scenario Generator scenario.

The need to generate reasonable yield projections and the shortfall of using traditional mean-variance models to perform the required simulations revealed the need for a new type of model. More advanced models would be required to directly model realistic economic variables and then use these economic variable projections to generate capital market returns that were both reasonable and consistent with historical return characteristics.

2.3.2 FINANCIAL TURMOIL IN THE 1970s DRIVES THE IMPERATIVE FOR FURTHER INNOVATION

The collapse of the Bretton-Woods system of fixed international exchange rates in the early 1970s led to a period of financial turmoil, as the U.S. dollar became the basis for international exchange and other currencies were allowed to float against the dollar. This engendered a period of financial turmoil, including stock market volatility, as the world adapted to new international rules.

The 1973–1974 stock market crash was the catalyst for the Report of the Maturity Guarantees Working Party (1980), a joint initiative of the Institute of Actuaries and the Faculty of Actuaries. The report focused on devising a reserving method related to maturity guarantees that would inspire a reasonable degree of confidence that insolvency would not result from the granting of maturity guarantees. This issue came into focus as a result of the policy value losses following the significant U.K. stock market declines that occurred during the 1970s.

A. D. Wilkie was the first to tackle this challenge as part of the modeling that was used to inform the Report of the Maturity Guarantees Working Party, and in 1986 he officially introduced a new type of economic and capital market

model that became known as the Wilkie model. This new model utilized stochastic regression techniques within a cascade model structure to generate simulations of both economic variables (interest and inflation rates) and capital market returns (stocks and bonds).

Over the years, the use of stochastic regression techniques that was the basis of the model Wilkie initially popularized has been overtaken by stochastic differential equation models informed by the research work within the field of option-pricing theory. Option-pricing models proved to be more theoretically precise than regression-based models and allowed for both real-world projections and the pricing of contingent cash flows, both of which are critically important for a large subset of financial modeling applications.

These new stochastic economic models were referred to by different names over the years, but the financial community eventually came to know them as economic scenario generators.

2.4 ESG MODEL APPLICATIONS

Later chapters will go into more detail about applications that are driven by economic scenario generators. However, it is useful to touch on the breadth and depth of these applications to get an appropriate appreciation of the growing importance of ESG models within the financial community.

Simulation modeling has become the standard for the planning, valuation and assessment of risk-taking activities within the insurance, pension and banking industries. Regulatory and rating agencies have embraced these models as valuable supplemental informational frameworks in the determination of the solvency and financial strength of financial institutions. Companies are using ESGs to drive financial modeling applications used to assist with the following work:

- **Asset liability management (ALM).** The evaluation and structuring of matching and strategic mismatch strategies, together with the determination of value-at-risk (VaR) metrics to measure and monitor ALM risk exposures.
- **Financial planning.** Evaluation of the risks and rewards associated with strategic business plans and alternative strategic initiatives. This includes merger and acquisition analysis, business growth strategies and capital management.
- **Regulatory compliance.** Use of internal risk management models in the satisfaction of global regulatory capital and solvency requirements.
- **Investment strategy.** Evaluation of the financial risk and reward opportunities associated with alternative asset allocation strategies involving fixed income, equity and alternative investments in light of internal liability obligations and the external economic and capital market environment.
- **Pricing.** Determination of the price of a set of contingent liability cash flows and evaluation of the consequences of alternative product pricing strategies under variable and uncertain economic and capital market conditions.
- **Reserving.** Setting appropriate reserve levels, given the complex distributions of uncertainty in the liabilities cash flows and economic environment.
- **Hedging.** Evaluation of the cost versus risk reduction benefits of hedging liability and capital market exposures.
- **Reinsurance structuring.** Evaluation of the benefits of alternative reinsurance structures.
- **Risk attribution.** Setting enterprise-wide risk budgets and monitoring the cumulative exposure to individual risk factors such as credit, equity and currency.

These are just a few of the myriad of business applications facilitated by ESG models. The number of business applications continues to grow as companies become more familiar with the capabilities of these systems and are able to successfully implement them in their day-to-day operations.

2.5 HOW ESGS DIFFER FROM OTHER ECONOMIC MODELS

Many different types of economic models are used within the financial services industry. Some are designed to give insight into how the economy works. These can be considered structural models that attempt to reflect cause-and-effect relationships. An example is price elasticity, which suggests that the price of a good or service has a direct

impact on the demand for that same good or service. Thus, increasing the price (cause) will reduce the demand (effect). These models are classified under the general category of econometric models.

Other types of economic models attempt to predict the future level of some key economic variable. These models seek to identify relationships between the movements of one set of economic variables (independent variables) and those of an economic variable of interest (dependent variable) for use in tactical decision making. For instance, an economist may attempt to predict the impact on inflation given information about changing GDP and unemployment levels. These models are classified under the general category of forecasting models.

Still other economic models, such as stochastic variable term structure models, are built specifically to determine the price of complex financial derivatives and insurance contracts with embedded options. These economic pricing models follow strict mathematical properties that are designed to satisfy risk-neutral and arbitrage-free conditions. Term structure models have been developing rapidly since the introduction of option-pricing theory in the early part of the 1970s.

Economic scenario generators are different from these other types of economic models, although they often utilize similar mathematical concepts and model equations. Economic scenario generators are designed to project economic and capital market variables over some specified future time period, for use in financial risk management applications. The purpose of an ESG is thus to provide a reasonable representation of the full range of future dynamics and distributions of economic and capital market variables, so that risk can be appropriately assessed and risk mitigation strategies can be evaluated. Therefore, ESG models are not intended to be predictive, and there is no requirement that the models provide any insight about why the economy works the way it does. The primary objective of an ESG is to produce a set of scenarios that represent the type of economic conditions that may occur in the future.

Economic scenario generators are often parameterized to reflect “real-world behavior.” That is, the scenario sets produced by an ESG are intended to best reflect what might actually happen in the future. To the extent that the analysis being performed requires the determination of prices of complex financial derivatives and insurance contracts with embedded options, the ESG needs to be able to produce scenario sets that can be used to calculate such prices. As noted previously, models used by the financial services industry to estimate prices follow strict mathematical properties designed to satisfy risk-neutral and arbitrage-free conditions. Thus, ESG models used to drive financial risk management applications that require pricing need to have market-consistent modeling capabilities.

Typical pricing problems are concerned with determining the price of a set of contingent cash flows as of the current date. Pricing becomes much more complicated when prices of contingent cash flows also need to be calculated at future points in time. The complication arises because future economic conditions are unknown, so the pricing needs to be done under multiple future uncertain states. Whereas the current price requires the development of one risk-neutral economic projection, the calculation of prices at future time periods requires thousands of different risk-neutral economic projections. This process is referred to as nested stochastic analysis and is computationally intensive.

2.6 LIMITATIONS OF ESG APPROACHES

Despite their advantages, ESG models continue to have some limitations. These involve model risk, sample error, processing time, challenges of reaching convergence, the amount of data and degree of expertise required, and difficulty in explaining the models to the people applying the data.

2.6.1 MODEL RISK

One of the biggest concerns with simulation modeling is model risk. Raw historical data and quantitative techniques alone cannot be relied on to generate a robust economic scenario generator. First, financial data has inherent limitations, such as inconsistent data series, differences in measurement and collection methods, and bad data. Second, financial history is filled with instances of unique and mitigating circumstances on which it would be unreasonable to project future trends. Third, there is a wide array of modeling techniques and analytical tools with different attributes, capabilities and precision levels. Finally, it is not practical to model every variable and every

outcome precisely, so the model developer must determine which attributes are the most important to capture. Consequently, vastly different simulation results are possible, depending on the ultimate choice of models and calibrations. Thus, the model user must take care to perform appropriate model validation and to understand the limitations as well as the strengths of the models being used for a particular application.

2.6.2 SAMPLE ERROR, REQUIRING EXTENSIVE SENSITIVITY ANALYSIS

The use of a simulation approach for modeling institutional liabilities is subject to its own set of problems. Since simulation only approximates the underlying distribution, it introduces the problem of sample error into the analysis. Thus, the model user is faced with the question of how close the simulated results come to the theoretically correct answer. Typically, this question cannot be answered through mathematical formulas, leaving the modeler to perform sensitivity analysis where multiple sample sets are run using a fixed number of simulations (refer to Tables 2.2 and 2.3). The variation around the statistic of interest is calculated to see if the results fall within an acceptable tolerance level. If the variation is too large, additional simulations must be added to the sample set. This process is repeated until enough simulations are included in the sample set to assure convergence within an acceptable tolerance level.

2.6.3 COMPLEXITY, DEMANDING UNACCEPTABLE PROCESSING TIME

Depending on the complexity of the model, the statistic of interest and the tolerance level, the number of simulations required for convergence could be extremely large. To reduce the number of required simulations and speed up the processing time, the model developer often builds variance reduction techniques into the simulation model. Variance reduction techniques are designed to reduce the number of simulations necessary for convergence. Examples of variance reduction techniques include the use of stratified sampling, antithetic variates and importance sampling. These techniques may not be effective if the problem involves a function of several random variables or when information related to the entire distribution or individual percentiles is desired.

2.6.4 CHALLENGES OF REACHING CONVERGENCE

The use of low-discrepancy sequences also has shown promising results in reducing the number of scenarios required for convergence. Low-discrepancy sequences, or quasi-random sequences, use mathematically generated deterministic samples instead of random samples to speed up the rate of convergence. These samples are generated to more evenly cover the area of the distribution being sampled. While low-discrepancy sequences allow for faster convergence to the mean and variance than Monte Carlo sampling, they have been generally limited to low-dimension problem spaces.

2.6.5 NEED FOR SUBSTANTIAL AMOUNTS OF DATA AND TECHNICAL EXPERTISE

Simulation models require advanced information systems and technical expertise. They are computationally intensive, often not user-friendly and may require more data and expertise than other liability analysis systems. While the advancement of computing technology has lowered the barriers associated with the computational intensity of simulation models, the knowledge and technical expertise of the audience for whom these models are designed has not necessarily kept pace.

2.6.6 POTENTIAL TO BE SEEN AS A BLACK-BOX SOLUTION

In addition to the mechanical difficulties of simulation, the results of simulation models often are perceived by the recipient to be mysterious. The term *black box* is often used in conjunction with such models. A black box refers to a model whose mathematical underpinnings are not disclosed. Unfortunately, the use of this term has often been extended to models that are sufficiently complex and poorly understood. Since simulation is typically used in cases where straightforward analytical approaches are not feasible, these uses will tend to lend themselves to complex modeling and thus to such criticism. It is imperative that the model developer communicate the model structure and present the model results in a way that the recipient of the analysis clearly understands.

2.7 CHAPTER SUMMARY

1. Recent advances in technology and simulation modeling now allow for more rigorous analysis where the intricacies of analytical methods would otherwise have restricted the depth and breadth of the analysis.
2. With proper care, simulation models can be the basis of an effective tool enabling institutional managers to gain a better understanding of the risk they are facing and thereby improve their strategic decisions.
3. Stochastic simulations are used to obtain numerical results in applications where closed-form analytical solutions are not readily available.
4. Stochastic simulation is used extensively in the area of financial risk management where both economic and capital market return projections are required.
5. ESG models are used to drive these stochastic simulation applications, with one of the earliest applications being introduced by A. D. Wilkie in 1986 as a result of modeling done in support of the Report of the Maturity Guarantees Working Party.
6. Applications that utilize ESG technology can be found within the insurance, banking and pension industries. These applications include generation of pro forma financial statement simulations for use in financial risk management, regulatory compliance, investment strategy, reinsurance structuring, pricing and hedging.
7. ESGs, while similar in form to other types of economic generation models such as pricing, forecasting and econometric models, differ from these models in that their primary goal is to generate a set of future economic scenarios that spans the range of possible future outcomes for use in risk management applications.
8. While economic scenario generators are extremely useful for gaining insight into future financial risk and rewards, like any model, they have limitations. Modeling the future dynamics of the economy and financial markets presents many challenges, such as accounting for extreme events and regime changes. It is imperative that users of these models understand the strengths and weaknesses of any particular ESG, so they can make sure that the ESG being used is appropriate for the analysis being performed.

2.8 CHAPTER REFERENCES

- Hull, John. 2006. *Options, Futures, and Other Derivative Securities*. Englewood Cliffs, N.J.: Prentice Hall.
- Markowitz, H. M. 1959. *Portfolio Selection: Efficient Diversification of Investments*. New York: John Wiley & Sons.
- Maturity Guarantees Working Party (Ford et al.). 1980. Report of the Maturity Guarantees Working Party. *Journal of the Institute of Actuaries* 107: 101–212.
- Panjer, H. H., ed. 1998. *Financial Economics with Applications to Investments, Insurance and Pensions*, The Actuarial Foundation.
- Wilkie, A. D. 1986. A Stochastic Investment Model for Actuarial Use. *Transactions of the Faculty of Actuaries* 39: 341–403.

Chapter 3: The Role of ESGs in Satisfying Regulatory Requirements

Regulatory concerns in insurance have intensified in recent years, fueled by the financial markets turmoil brought about by the financial crisis of 2008, but also as a result of greater product complexity that has evolved in many parts of the insurance industry. In many cases, these regulatory concerns require a response that demands a stochastic view of future developments in both financial markets and the economy. Regulation requires consistency across entities such that regulators can both understand an entity's approach to risk and adequately benchmark solvency concerns against those risks. This need for a regulator to have comparability across entities can lead to conflicts with the desires and uses of a company's own internal risk models.

This chapter will focus on the role that ESGs play in responding to needs of the regulatory environment. Depending on the regulatory jurisdiction, regulatory requirements may be prescriptive (rules based) or may outline specific or general principles (principles based).

Section 3.1 will describe how the use of ESG models are increasingly satisfying evolving regulatory requirements in areas of solvency, as well as areas of pricing, reserving and capitalization. In most frameworks, the use of ESGs is included or considered in possible responses to evolving regulatory rules and principles, but the development of strict definitions or requirements for ESG applications is still emerging. In the United States under the National Association of Insurance Commissioners guidelines, ESGs are only beginning to emerge as an interesting, potentially preferred solution to solvency and product issues. ESGs have made the most headway in life insurance reserving regulations at the product level. In emerging Solvency II regulations and prescriptions, standard and internal models used for quantifying risk and establishing minimum capital requirements may make use of ESGs.

Section 3.2 addresses specific regulatory requirements with respect to ESGs in terms of calibration, validation and documentation. The NAIC has its own ESGs applicable for interest-sensitive and variable market-sensitive products in life and annuities. These ESGs develop only a few economic factors. Insurers may have other significant economic factors, such as currency exchange rates, that would need to be consistent with the regulatory-specified ESG. Solvency II outlines considerations that affect the selection and use of an ESG. These include extrapolation beyond last liquid tenor progressing to an ultimate forward rate; matching adjustment; symmetric adjustment; and volatility adjustment.

Section 3.3 discusses general ESG considerations that need to be taken into account when using ESGs for regulatory reporting purposes. These include time horizon and risk standards, governance, calibration standards, consistency with unrelated entities, potential conflicts with internal models, principles-based capital calculations, and the level of expertise required to reconcile internal practices with regulatory requirements.

3.1 APPLICATIONS OF ESGS IN RESPONDING TO EVOLVING REGULATORY REQUIREMENTS

ESGs are used in regulatory applications in two distinct ways. First, they are used to establish a time-zero value for a regulated entity's assets, liabilities and surplus, according to the valuation principles required by the corresponding regulator. Second, they are used to simulate the real-world evolution over time of the entity's assets, liabilities and surplus. These regulatory applications are used to test a company's solvency and liquidity position as well as in support of individual product line pricing and reserving. The first use establishes a value of the available surplus assets at the valuation date (time zero), which can be compared with a benchmark for required surplus assets to absorb future realization of the entity's risk exposures. The second is used to develop the standard(s) for required surplus assets (including liability reserves) by quantifying the variations due to the realization of a company's risk exposures.

ESGs are being used in multiple regulatory frameworks across the globe. In most frameworks, the use of ESGs is included or considered in a possible response to evolving regulatory rules and principles, but the development of strict definitions or requirements for ESG applications is still emerging. The most advanced frameworks are seen in the emerging principles of ComFrame (a common framework for international insurance supervision established by the International Association of Insurance Supervisors) and in developing rules in EIOPA (the European body representing the regulatory authorities of the European Union and its developing solvency rules embodied in Solvency II).

3.1.1 INTERNATIONAL ASSOCIATION OF INSURANCE SUPERVISORS (IAIS): COMFRAME AND INSURANCE CORE PRINCIPLES (ICPs)

The International Association of Insurance Supervisors is an organization that has no regulatory power itself but brings together insurance regulators with such power (as well as interested parties) to try to coordinate on international insurance regulatory issues.

In the aftermath of the global financial crisis of 2008, regulators in the financial sphere became acutely aware of the danger of systemic risk. As a result, the IAIS developed a Common Framework (ComFrame):

The Common Framework for the Supervision of Internationally Active Insurance Groups (ComFrame) is a set of international supervisory requirements focusing on the effective group-wide supervision of internationally active insurance groups (IAIGs). ComFrame is built and expands upon the high level requirements and guidance currently set out in the IAIS Insurance Core Principles (ICPs), which generally apply on both a legal entity and group-wide level.¹

The ICPs on which the ComFrame is built include implicit references on the use of ESGs (emphasis in **bold** is ours):

16.1 The supervisor requires the insurer's enterprise risk management framework to provide for the identification and quantification of risk **under a sufficiently wide range of outcomes using techniques which are appropriate to the nature, scale and complexity of the risks the insurer bears** and adequate for risk and capital management and for solvency purposes.

16.1.14 Stress testing measures the financial impact of stressing one or relatively few factors affecting the insurer. Scenario analysis considers the impact of a combination of circumstances which may reflect extreme historical scenarios which are analyzed in the light of current conditions. **Scenario analysis may be conducted** deterministically using a range of specified scenarios or **stochastically, using models to simulate many possible scenarios, to derive statistical distributions of the results.**²

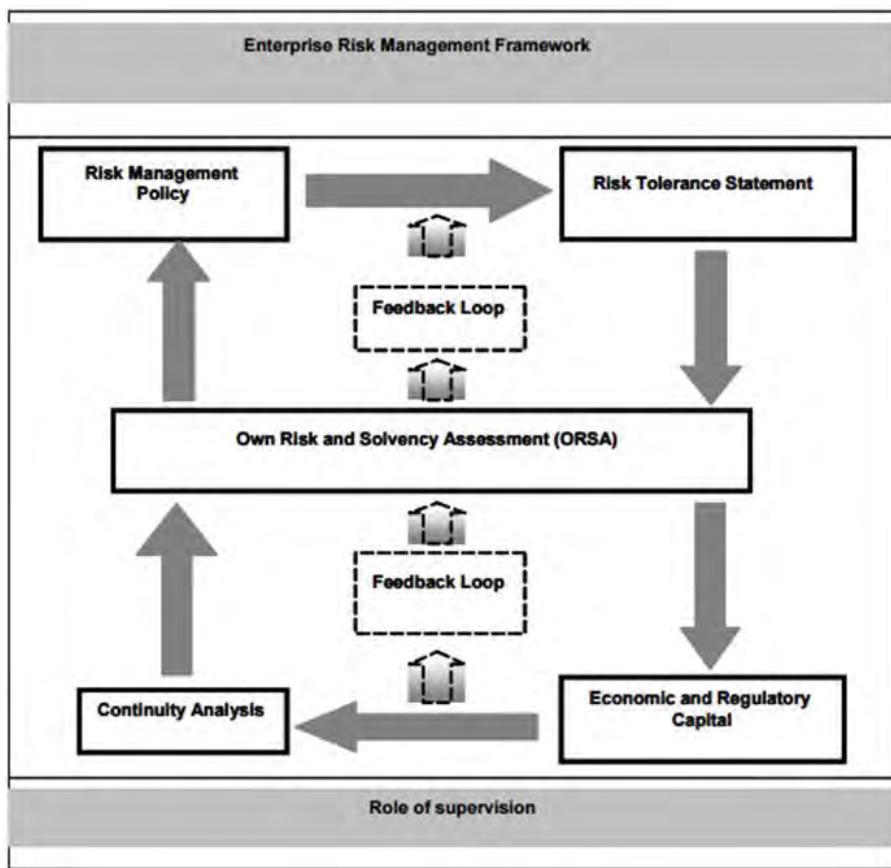
In these references, allowance is made for stochastic analysis that achieves a sufficiently wide range of outcomes, reflects the scale and complexity of risks, and provides the ability to determine a statistical distribution of results. ESGs are well equipped to achieve these objectives, but for now the guidelines permit, but do not require, this level of analysis.

Developing a range of scenarios, including economic scenarios, is described in 16.1, which deals with risk identification and measurement. ICP 16 as a whole deals with an enterprise risk management framework, as illustrated in Figure 3.1.

¹ International Association of Insurance Supervisors (IAIS), "Common Framework," <http://iaisweb.org/index.cfm?event=getPage&nodeld=25229>, accessed January 11, 2016.

² IAIS, "Insurance Core Principles," updated November 2015, <http://iaisweb.org/index.cfm?event=getPage&nodeld=25224>.

Figure 3.1
THE IAIS STANDARD ENTERPRISE RISK MANAGEMENT FRAMEWORK



Source: IAIS, Insurance Core Principles, updated November 2015, p. 168.

ESGs can be used in multiple parts of this ERM framework, specifically in the calculation of economic and regulatory capital, continuity analysis and the ORSA. ICP 16 addresses multiple areas on modeling the level of risk and identifying relationships between risks. The ICPs are high-level principles and do not specify quantitative requirements that the actual insurance supervisors require insurers to follow. These elements are embodied in specific regulatory systems inasmuch as they pertain to ESGs.

3.1.2 USE OF ESGS IN U.S. REGULATORY CAPITAL AND RESERVING REQUIREMENTS

In the United States under the National Association of Insurance Commissioners guidelines, ESGs are only beginning to emerge as an interesting, potentially preferred solution to solvency and product issues. ESGs have made the most headway in life insurance reserving regulations at the product level, and they are beginning to be recognized as an alternative in emerging requirements for own risk solvency assessment (ORSA). NAIC guidance states that individual state insurance commissioners may provide input on the level of the measurement metric to use in the stressed condition or specify particular parameters used in the economic scenario generator.

The following timeline in Table 3.1 shows the history and evolution of capital requirements in the United States.

Table 3.1

TIME LINE FOR CAPITAL REQUIREMENTS, NAIC

1970s	Concept of risk-based capital (RBC) formulas developed by Society of Actuaries
1990	NAIC begins investigation of RBC
1993	RBC for life in effect
1994	RBC for property and casualty in effect
2000	C3 Phase I: scenario testing for single-premium life and fixed annuities
2005	C3 Phase II: principles-based RBC for variable annuities with guarantees
2008	SMI announced
2009	AG VACARVM: principles-based reserving for variable annuities
2012–2015	Review of bond default factors (C1/R1); will recalibrate

The core of capital requirements in the U.S. system is risk-based capital (RBC). First implemented in the 1990s, RBC was intended as a tool for regulators to determine which companies were weakly capitalized in order to better focus their supervisory activities. The current risk-based approach has incremental capital positions for regulatory intervention: two as “action” levels and two as “control” levels.

RBC has had fairly simple core equations, one for life companies and another for nonlife. On the nonlife side, there has not been any significant use of ESGs in calculating RBC.

On the life and annuity side, there has been an increasing regulatory requirement for the use of ESGs, but currently they apply at a product level with respect to variable products with embedded financial guarantees, as opposed to overall balance sheet considerations. There, a simple factor approach had been seen as possibly underestimating the risks in some products. The first project to update the risk charge for this, primarily affecting fixed annuities and single-premium life products, was called C3 Phase I. C3 Phase I developed the asset-liability mismatch risk charge by having asset-liability models for the covered lines of business through a relatively small set of scenarios (12 or 50) generated from a standard interest rate ESG model. This interest rate ESG model was calibrated to produce real-world scenarios, and the subsets were chosen based on testing conducted by the American Academy of Actuaries. The model used for this purpose was a two-factor model, which was double mean-reverting.³

The next such project, C3 Phase II, covered variable annuities and their guarantees. It was much more complicated than C3 Phase I. C3 Phase II involved the use of 10,000 ESG-generated scenarios and added equity market models in addition to the C3 Phase I interest rate model. The equity market model was a stochastic log volatility model,⁴ with correlations between certain major equity indices. In addition to stochastic modeling, there was a Standard Scenario intended to provide a minimum deterministic capital requirement. This was followed by a related project, AG Commissioners Annuity Reserve Valuation Method for Variable Annuities (VACARVM), which used the same economic models to develop principle-based reserves for variable annuities.

There was an update to the interest rate model in 2010 that moved from a fixed mean reversion point (MRP) to an annually changing one. The new model has the same structural form of a double-mean-reverting, real-world interest rate model for the United States, with a new parameterization. As of January 2016, this updated model has not been approved for use for life RBC but has been approved for use in the new valuation manual with principles-based reserving for life insurance: VM-20.

³ A double mean-reverting model is a model in which two of its stochastic factors have a mean-reversion component, and the two factors generally interact.

⁴ A stochastic log volatility model is an equity model in which the stock index price has the standard lognormal stochastic equation but the natural logarithm of the volatility also is driven by a stochastic process (as opposed to being a constant).

3.1.3 UNITED STATES: NAIC ORSA

While RBC has made limited use of ESGs, ORSA, as part of an ERM framework seen in the ICPs and ComFrame, has no explicit ESG restrictions. Ultimately this is under the control of the individual state regulators, because insurance regulation is primarily established on a state basis, though there is coordination with the NAIC.

The *NAIC ORSA Guidance Manual* states (as of July 2014), “The commissioner may provide input to an insurer’s management on the assumptions and scenarios to be used in its assessment techniques. For assumptions that are stochastically modeled, the commissioner may provide input on the level of the measurement metric to use in the stressed condition or specify particular parameters used in the economic scenario generator.”⁵

As the first required ORSA reports were submitted in 2015, we do not yet have information as to whether specific state commissioners have provided any guidance on the ESG choice and parameterization. The ORSA manual itself is a high-level document with few prescriptions, but the amount of prescription is likely to come out in actual practice.

3.1.4 EUROPEAN UNION: SOLVENCY II

In emerging Solvency II regulations and prescriptions, standard and internal models used for quantifying risk and establishing minimum capital requirements may make use of ESGs. ORSA requirements are also specified in Solvency II, which may be performed using a model that includes an ESG. There is an internal model approval process prescribing governance procedures and a use test. However, the application of the model, in particular the ESG, to other management decisions does not have to be based on the same calibration as that used for the regulatory capital application.

The European Insurance and Occupational Pensions Authority (EIOPA), is the body representing the national regulators of the EU member states. It has an advisory role to the EU bodies legally charged with developing EU solvency rules (the European Commission, the European Parliament, and the Council of Ministers). It also has some delegated powers to implement the regulations in a consistent manner across all EU member states.

EIOPA is not a regulator itself. Rather, it is a body that represents and coordinates the perspectives and activities of EU member states’ regulators. A regulated entity therefore has direct dealings not with EIOPA itself but with the regulator of the entity’s home country. In line with the variations in size and market traditions among the various EU member states, there is a corresponding variation in the size and experience of the staff of the member states’ regulators. This can lead to variations in practice to a somewhat greater extent than would be expected in the case of a market with a single regulatory body.

Solvency II is the regulatory regime being introduced across the EU with effect from January 1, 2016. Solvency II consists of three “pillars”: quantification of required capital, regulatory supervision process, and disclosure.

ESGs are affected by all of these three pillars. They play a role in the quantification of risk and required capital through their use in full or partial internal models. Their use is subject to the governance and process framework in pillar 2. And they, as well as their use, are subject to the disclosure requirements of pillar 3.

The solvency capital requirement (SCR) and minimum capital requirement (MCR) are different thresholds of regulatory required capital and trigger different levels of regulatory and/or company action. These may be assessed using an entity’s full or partial internal model, which may include the use of an ESG.

The ORSA is an own risk and solvency assessment that entities are required to perform periodically. This may also be performed using a model that includes an ESG. However, since this is an own risk and solvency assessment, the ESG does not need to be calibrated to the ESG used within the internal model for the SCR or MCR.

For companies that do not use an internal model, a standard formula is available. Entities that use the standard formula may be subject to a capital add-on if the regulator judges that the assumptions underlying the standard

⁵ National Association of Insurance Commissioners, *NAIC Own Risk and Solvency Assessment (ORSA) Guidance Manual*, as of July 2014, pp. 7–8, accessed at <http://www.naic.org/store/free/ORSManual.pdf>.

formula are not reflected in the entity's business or operations. Entities must therefore be able to demonstrate whether the standard formula is applicable. The standard formula may be modified by using specific parameters.

For companies that use an internal model, there is an internal model approval process (IMAP). Internal models require the approval of the regulator before they can be used for regulatory capital assessment. The IMAP requires, among other things, the existence of a governance process including a model change policy and the demonstration of the management's confidence in the internal model through a "use test." The model must be used for management decisions other than solely regulatory capital. However, the application of the model—in particular, the ESG—to other management decisions does not have to be based on the same calibration as that used for the regulatory capital application.

Entities are free to use any ESG they wish as part of their internal model or ORSA, subject to the IMAP requirements. In practice, the IMAP scrutiny of the ESG will usually include a comparison with the assumptions in the market risk component of the standard formula.

3.1.5 SWITZERLAND (FINMA): SWISS SOLVENCY TEST (SST)

FINMA is the regulatory authority for financial services firms in Switzerland. Swiss insurers are subject to the Swiss Solvency Test (SST) rather than the EU Solvency II. SST has been in place for several years and shares many of the same principles as Solvency II. There are differences in the detailed requirements and in the approval process. Switzerland has received recognition by the EU that its solvency regime provides protection for policyholders comparable to that of Solvency II. This is called "equivalence" and means, among other things, that EU entities can rely on the SST of their Swiss subsidiaries and reinsurers without needing them to be Solvency II compliant themselves. Like Solvency II, SST has a standard formula for entities that are not required to use an internal model.

With regard to ESGs, the SST includes the requirement for entities to report on the impact of a set of extreme economic scenarios and sensitivities of market risk factors that are specified by FINMA.

3.2 SPECIFIC REGULATORY REQUIREMENTS WITH RESPECT TO ESGS

The NAIC, Solvency II and the Swiss Solvency Test each have specific regulatory requirements that apply to use of ESGs.

3.2.1 UNITED STATES: NAIC

The NAIC has its own ESGs that are primarily applicable to regulatory considerations for interest-sensitive and variable market-sensitive products in life and annuities. These ESGs develop only a few economic factors: U.S. government yield curves and a handful of equity indices. Insurers may have other significant economic factors, such as currency exchange rates, that would need to be developed consistent with the regulatory-specified ESG. If missing economic factors are correlated with those generated by the NAIC ESG, then this can create difficulties in producing full scenario sets.⁶

3.2.2 EUROPEAN UNION: SOLVENCY II

Solvency II outlines several considerations that affect the selection and use of an ESG. These include the following four:

1. Extrapolation beyond last liquid tenor, progressing to an ultimate forward rate
2. Matching adjustment
3. Symmetric adjustment
4. Volatility adjustment

⁶ American Academy of Actuaries generator hosted at Society of Actuaries website, <https://www.soa.org/research/software-tools/research-scenario.aspx>, accessed January 11, 2016.

The Solvency II requirements for the yield curves from an ESG include the need to extrapolate from the last liquid tenor. The proposed methodology involves an extrapolation curve that smoothly progresses to an “ultimate forward rate” set by the regulator, which may vary from country to country.

The matching adjustment is the adjustment of the discount rate for the liabilities to reflect, to the extent permitted, the entity’s investment in illiquid matching assets.

The symmetric adjustment is an adjustment mechanism to the equity index level in the market risk module of the standard formula for the solvency capital requirement. The stated purpose is to “mitigate undue potential pro-cyclical effects of the financial system and avoid a situation in which insurance and reinsurance undertakings are unduly forced to raise additional capital or sell their investments as a result of unsustainable adverse movements in financial markets.”

The volatility adjustment is another mechanism intended to prevent pro-cyclical investment behavior. With the volatility adjustment, “insurance and reinsurance undertakings may be allowed to adjust the relevant risk-free interest rate term structure for the calculation of the best estimate of technical provisions to mitigate the effect of exaggerations of bond spreads.” The volatility adjustment is “based on reference portfolios for the relevant currencies of those undertakings and, where necessary to ensure representativeness, on reference portfolios for national insurance markets.” To ensure adequate transparency, companies are required to publicly disclose the impact of the volatility adjustment on their financial position.

A useful reference summarizing additional details for the so-called Long Term Guarantee Package, which includes the matching and volatility adjustments, can be found in McHugh and Schiffel (2014), a paper titled “Omnibus II Agreement on Long-Term Guarantee Package and Transitional Measures.”

3.2.3 SWITZERLAND: SWISS SOLVENCY TEST

The Swiss Solvency Test (SST) requires that the output of an entity’s internal model be supplemented by the impact of a set of extreme economic scenarios and sensitivities of market risk factors specified by the regulator.⁷

3.3 GENERAL REGULATORY ESG CONSIDERATIONS

Considerations relating to the application of ESGs for regulatory purposes may be considered across all jurisdictions. This section discusses seven such factors:

1. Time horizon and risk standards
2. Governance
3. Calibration
4. Consistency across unrelated entities
5. Conflicts with internal models
6. Principles-based capital calculations
7. Expertise

3.3.1 TIME HORIZON AND RISK STANDARDS

The time horizon and risk standards for regulatory purposes are set by the regulator. While the regulated entity may use other horizons and/or risk standards for its own strategic and risk management purposes, it will need to conform to the regulator’s parameters for regulatory purposes. A typical regulatory time horizon is one year, while the entity’s time horizon could be longer, depending on the nature of its business and markets and on its internal planning and reporting processes. Risk standards vary by entity, as they need to be responsive not only to the regulators but also to other stakeholders such as policyholders, rating agencies and shareholders.

⁷ FINMA “Swiss Solvency Test (SST),” <https://www.finma.ch/en/supervision/insurers/cross-sectoral-tools/swiss-solvency-test-sst/>, accessed January 15, 2016.

3.3.2 GOVERNANCE

For regulatory applications, ESGs are subject to an approval process by the regulators. The approval process requires that management have an acceptable governance process in place, including a model change policy, and that senior management and board members have an adequate level of understanding and appreciation of the ESG—in particular, its relevance and limitations.

The documentation and disclosure of ESGs, whether internally developed or purchased from a specialist vendor, must include a description of the underlying models, the reasons for choosing these models, the source data from which the models are calibrated, the methodology for fitting model parameters to the data, any expert judgments applied, and the validation of the resulting calibrated models. This disclosure is not public disclosure, but rather refers to the disclosure to the regulatory body.

There are increasing regulatory requirements related to each of the following:

- Appropriate qualifications and experience to be met by holders of key functions, as well as fit and proper requirements.
- Understanding by senior executives and board members of the models used by entities in their regulatory solvency assessments. This especially applies to an understanding of the limitations of the models.
- Openness and transparency of the internal part of the regulatory solvency assessment.

ESGs used within the regulatory capital assessment process therefore not only have to be technically sound; they also need to conform to the management/regulatory process in which the ESG is embedded and the governance framework around that. This means that ESGs are subject to increasing requirements for documentation and disclosure.

3.3.3 CALIBRATION

One key issue in model governance is that of controls regarding the calibration and validation of the internal models. When internal models are created and used, it is expected that the parameterization of the model will vary over time. Regulatory requirements may be less flexible to the extent that they are prescriptive, and differences will need to be reconciled and disclosed.

3.3.4 CONSISTENCY ACROSS UNRELATED ENTITIES

Consistency across entities and across time is important to regulators. This includes consistency across owned but unrelated entities, so regulators can adequately compare results. Consistency with other, unrelated entities is perhaps less important from the perspective of a particular entity conducting its own internal risk management or shareholder reporting activities. Therefore in regulatory applications, ESGs will be subject to consistency requirements. As well, regulated entities will often find they need to use different choices in ESGs, at least with respect to calibration, to perform nonregulatory business activities.

Regulators may impose constraints on the ESG in order to achieve consistency with their particular objectives. Sometimes these constraints may seem arbitrary or at odds with economic principles of asset and liability valuation but are put in place to satisfy a regulatory goal. These regulatory constraints may be set up to provide easier comparability of regulated entities or to prevent these entities from “gaming” the results.

3.3.5 CONFLICTS WITH INTERNAL MODELS

On occasion, regulatory applications may result in unintended incentives to companies—that is, incentives for companies to operate in a manner that goes against sound risk management principles. Regulatory requirements, which apply to the entire market, therefore have to be set with care to avoid generating systemic risks. For example, certain asset-liability regulatory valuation structures might inadvertently lead companies to invest pro-cyclically—that is, to buy when prices are high and sell when prices are low.

For example, a market risk-free (government) yield curve is a typical starting point for valuation of projected cash flows. However, the market may be relatively illiquid at longer maturities, although there may be significant liability

cash flows expected at these long maturities. Different ESGs may extrapolate yields differently beyond the last liquid point, which could have a significant effect on valuations if an entity switched from one ESG to another. Therefore, regulators may impose extrapolation criteria and/or specific methodologies and parameters for carrying out the extrapolation process.

These regulatory restrictions can be somewhat controversial, since they are to some extent subjective and (therefore) not universally agreed upon. These controversies can occur not only between regulators and the regulated entities but also between differing regulators themselves.

A further divergence, or at least point of controversy, between regulatory and internal usage of an ESG results from the illiquidity of liabilities. Entities may exploit in their investment strategy the fact that many of their liabilities are illiquid and will not need to be realized at an arbitrary date. This allows them to benefit from investing in relatively illiquid assets that provide a spread to liquid asset alternatives. A point of contention is about whether such an entity may or may not decrease the value of its liabilities in line with the higher yields being generated by the illiquid assets it is holding to match the liabilities. While the entities themselves may wish to recognize such an “illiquidity premium,” regulators may not allow them to do so, in order to provide more protection to policyholders.

3.3.6 PRINCIPLES-BASED CAPITAL CALCULATIONS

The regulatory trend has been toward market-consistent valuation principles. A second trend has been toward a custom approach to risk and capital assessment and away from a standard, one-size-fits-all approach. A third trend, driven by the first two trends, has been toward a principles-based approach rather than a prescriptive rules-based approach.

However, rules-based approaches still apply to some key aspects for which there may be wide deviations possible in the face of limited objective evidence. These rules-based approaches ensure consistency and avoid entities selecting unduly favorable assumptions.

3.3.7 EXPERTISE ISSUES

As a result of the mathematics of ESGs being relatively sophisticated, there is a shortage of skills and experience globally. Since regulators' budgets tend to be constrained to a greater extent than those of regulated entities, it is not surprising that regulators struggle to attract and retain individuals with appropriate ESG skills. Anecdotal evidence suggests that this has been a factor in insurers' different experiences in working with regulators to achieve internal model approval with respect to their ESG, which may be unrelated to the ESG itself. This situation will certainly evolve, but it is hard to predict whether any change will be positive or negative for regulated entities.

3.4 CHAPTER SUMMARY

1. ESGs are being used more frequently and in a more widespread manner in satisfying regulatory activities in the areas of pricing, reserving, capitalization and solvency.
2. Significant validation and documentation requirements govern the use of ESGs for regulatory reporting.
3. Different regulatory regimes provide differing levels of scrutiny and freedom with respect to ESG model choice and calibration.
4. There is a balancing of regulatory need for comparability across companies versus a company's own need to reflect specific risks in its own asset-and-liability portfolio.
5. ESG output used for regulatory purposes may not be appropriate to use to manage the business, because ESG model type and calibration may differ widely, reflecting differences in the goals and viewpoints of regulators and insurers.

3.5 CHAPTER REFERENCES

3.5.1 INTERNATIONAL

- International Association of Insurance Supervisors. 2016. Common Framework.
<http://iaisweb.org/index.cfm?event=getPage&nodeId=25229>. Accessed January 11, 2016.
- . 2015. Insurance Core Principles. November, <http://iaisweb.org/index.cfm?event=getPage&nodeId=25224#>.

3.5.2 UNITED STATES

- American Academy of Actuaries. 1999. Phase I Report of the American Academy of Actuaries' C-3 Subgroup of the Life Risk Based Capital Task Force to National Association of Insurance Commissioners' Risk Based Capital Work Group. October, p. 16. Accessed at https://www.actuary.org/files/C3_Phase_I_Report_October_1999.pdf.
- . 2005. Recommended Approach for Setting Regulatory Risk-Based Capital Requirements for Variable Annuities and Similar Products. Presented by the Life Capital Adequacy Subcommittee to the National Association of Insurance Commissioners' Capital Adequacy Task Force. June, pp. 25–34. Accessed at http://www.actuary.org/pdf/life/c3_june05.pdf.
- . 2016. Economic Scenario Generators. Accessed January 11, 2016, at <http://www.actuary.org/content/economic-scenario-generators>.
- National Association of Insurance Commissioners. 2014. *NAIC Own Risk and Solvency Assessment (ORSA) Guidance Manual*. As of July 2014, pp. 7–8. Accessed at <http://www.naic.org/store/free/ORSManual.pdf>.
- . 2016. Own Risk and Solvency Assessment (ORSA). Accessed January 11, 2016, at http://www.naic.org/cipr_topics/topic_own_risk_solvency_assessment.htm.
- Society of Actuaries. 2016. Economic Scenario Generators. Accessed January 11, 2016, at <https://www.soa.org/research/software-tools/research-scenario.aspx>.

3.5.3 SOLVENCY II

The Solvency II directive, the Omnibus II directive and the Delegated Acts provide the legal framework for the implementation of Solvency II. They can be found at the European Union sources in the following list.

- European Union. 2009. Directive 2009/138/EC of the European Parliament and of the Council of 25 November 2009 on the Taking-up and Pursuit of the Business of Insurance and Reinsurance (Solvency II). *Official Journal of the European Union* (December 17), <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32009L0138&from=EN>.
- . 2014. Directive 2014/51/EU of the European Parliament and of the Council of 16 April 2014 (Omnibus II directive). *Official Journal of the European Union* (May 22), <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32014L0051&from=EN>.
- . 2015. Commission Delegated Regulation (EU) 2015/35 of 10 October 2014 Supplementing Directive 009/138/EC of the European Parliament and of the Council on the Taking-up and Pursuit of the Business of Insurance and Reinsurance (Solvency II). *Official Journal of the European Union* (January 17), http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:JOL_2015_012_R_0001&from=EN.
- McHugh, Manijeh, and Simon Schiffel. 2014. *Omnibus II Agreement on Long-Term Guarantee Package and Transitional Measures*. Solvency Consulting Knowledge Series. Munich RE, November, accessed at http://www.munichre.com/site/corporate/get/documents_E48679993/mr/assetpool.shared/Documents/5_Touch_Publications/302-08521_en.pdf.

Chapter 4: Business Applications of ESGs in the Insurance and Pension Industries

This chapter will address practical applications and business uses of ESGs across life and P&C insurance companies, as well as for pension analysis. Economic scenario generators are often integrated into key planning and decision-making activities. The examples in this chapter are not meant to be exhaustive, but rather to highlight key uses of ESGs in a business setting. Some applications, such as strategic asset allocation, are relevant in all three areas. However, there are important aspects of ESGs that will differ among the three fields, or even within a field or product, due to different fundamental questions being addressed or different behavior of the associated liabilities.

Section 4.1 will describe the application of ESGs to address specific problems in life insurance. Applications in life insurance focus primarily on interest rate changes and ways policyholder behavior may dynamically affect liabilities. Some of the considerations explored in life insurance applications may also apply to pensions and property and casualty applications.

Section 4.2 describes the similarity of specific applications in pensions, focusing primarily on defined benefit pension approaches. In addition, considerations of the interaction of economic factors applying to pension sponsors and pension assets and liabilities, including dynamics of payroll and retirement decisions, also may be better understood through ESG modeling.

Section 4.3 describes applications in property and casualty insurance, which by the nature of its liabilities is quite different from life insurance and pensions. The impact of inflation, subject to ESG modeling, plays a greater role in understanding liabilities (loss reserves), and other economic factors such as GDP growth and employment play a greater role in pricing cycle analysis and enterprise risk management analysis.

4.1 LIFE INSURANCE APPLICATIONS

Applications of ESGs in life insurance are focused on the interaction of interest rate changes and policyholder behavior often connected to interest rate changes. Life insurance results at both product level and enterprise level are determined by the interaction of investment performance on assets (built up from the collection of premiums) and the payout of liabilities (based on events of mortality, morbidity (in health or disability), policy surrender influenced by crediting strategies, or annuity payout.) Because of the complexity of the interaction of these factors over an extended time horizon, an ESG provides a comprehensive tool to understand the range of potential outcomes.

An ESG can also produce a large number of scenarios, which can be conflated to a probability distribution, so the likelihood of certain events can be anticipated. This insight is critical to making informed judgments about the potential costs and rewards of the risks involved, allowing companies to progress from strict risk elimination to risk mitigation within the context of the cost of risk.

Several applications in life insurance are best understood through the use of an ESG. Considerations regarding the choice of market-consistent versus real-world scenarios, number of scenarios and forecast horizon, specific economic factors to model, and other features such as specific calibration approaches to apply may depend on the specific application being considered.

The following applications in life insurance are representative of those best understood through the use of an ESG:

- Life liability valuation
- Effective duration analysis
- Stress testing
- Economic capital
- Strategic asset allocation

4.1.1 LIFE LIABILITY VALUATION

Life insurance liability valuations require a focus on the interaction of interest-sensitive cash flows and policyholder behavior regarding lapses and other optionality. ESGs are necessary to properly respond to the challenges represented by this interaction. Both the potential range of scenarios, reflecting possible extreme events, and the likelihood of scenarios are important in developing a consistent valuation. In selecting appropriate ESG models for this application, market-consistent scenarios are required to properly understand the cash flows associated with the volatility of implied forward yield curves, and a large number of scenarios may be needed to properly represent value changes under extreme conditions.

A single scenario projection will not divulge the full spectrum of potential results that may occur when a life insurance company creates a long-term contract that allows policyholders the opportunity to act in their best interest to the potential detriment of the insurer. When interest rates move lower, minimum rate guarantees may be set by a crediting rate floor, which can become very expensive to an insurer, as the yield earned on their investment portfolio may not support the credited rate. Conversely, rates may move higher, which may allow for competing insurance products to appear more attractive to policyholders. This could lead to higher lapse rates at a time when fixed-income assets have fallen in value, leading to untimely asset liquidation of the insurer's portfolio. With many life insurance contracts, the insurer has essentially sold interest rate options on both sides of the balance sheet, which causes the insurer to lose value during times of large interest rate changes.

Therefore, a variety of scenarios are needed to cover the possible outcomes of the insurance contract, not simply a single scenario. An ESG is required in this case, and the scenarios produced must adequately account for the range and likelihood of plausible outcomes.

In addition to gaining insight into the full spectrum of potential results, understanding the distribution of these results over the range and likelihood of interest rate movements and policyholder responses is important in order to get a consistent valuation of the liabilities. The number of scenarios required to get a consistent valuation depends on the sophistication of the implied options sold and the extent to which the options change in value. If large value changes are expected during extreme scenario events, more scenarios are likely needed in order to populate enough of these wide-changing events for proper representation. Thousands or even hundreds of thousands of scenarios may be needed to fully represent the most extreme events that may occur.

As a practical matter, however, long model run times may prohibit the use of a very large number of scenarios, so a balance must be struck between accuracy and reasonable run times. Certain methods, such as scenario reduction techniques that would oversample or reweight the tail events, or model reduction techniques including liability proxy modeling, may increase model speed to enable the use of more scenarios.

The use of stochastic simulations for valuing liabilities in life insurance applications requires an ESG that produces market-consistent scenario sets. This means that the implied forward curves, on average across all scenarios, and the volatility of those curves are informed by actual prices of traded bonds and options in the market. Market-consistent scenarios are used for properly valuing financial instruments so that fixed or contingent projected cash flows can be consistently valued. Current capital market valuation sets the price of these cash flows, and market-consistent scenarios assure that all cash flow valuations and asset or liability flows have a consistent relative value. Market-consistent scenarios produce different dynamics than real-world scenarios, particularly for interest-sensitive cash flows, so it is important to use these appropriately.

4.1.2 EFFECTIVE ASSET AND LIABILITY DURATION ANALYSIS

The ability to better value asset or liability financial instruments through the use of an ESG allows for more precise effective-duration calculations. Effective duration measures the interest sensitivity of financial instruments by measuring the price change of a set of asset or liability cash flows for a given level of interest rate change. Liabilities that do not have interest-sensitive cash flows (that is, they do not have embedded interest rate options) may be valued using a single market-consistent interest rate projection that follows the forward curve. However, for liabilities (and assets) with interest-sensitive cash flows, a single scenario is not sufficient for valuation, so a set of market-consistent scenarios is needed. This set of scenarios can then be shocked up or down to produce the new liability values for use in the duration calculation:

$$\text{Effective duration} = \frac{V_{-\Delta y} - V_{+\Delta y}}{2(V_0)\Delta y},$$

where Δy is the change in interest rates and V is the resulting value of the liability given the interest rate change.

Typically this constant rate shock is applied to the implied spot curve, which has the effect of changing the forward curves for each set of shocked scenarios. Therefore, simply applying the shocks to all rates at all time periods is inconsistent, as they will not be risk neutral; a new set of market-consistent scenarios is needed for each changed set of scenarios.

Liability effective durations have traditionally been a useful metric for communicating the interest rate sensitivity of liabilities to investment departments.

4.1.3 STRESS TESTING AND CASH FLOW TESTING

ESGs are an essential companion to traditional actuarial software in performing deterministic stress testing and regulatory cash flow testing.

Life insurance companies are heavy users of actuarial software, through which it is possible and even routine to project company results into the future across a variety of potential economic and capital market states. Insurance product performance is affected by external economic and capital market events, including through secondary effects that are difficult to directly calculate. For example, policyholder behavior such as lapses can be influenced by whether interest rates are rising or falling, especially if the change is extreme. To anticipate the impacts of changing product performance on an insurer's financials, stochastic models of policyholder behavior are used to reflect these changes, as well as changes in asset and liability value. This is a key component of deterministic stress testing and a primary feature of regulatory cash flow testing that has been in use for decades to evaluate the adequacy of life insurance reserves.

In traditional deterministic stress testing, somewhat arbitrary adjustments are added or subtracted from the current rate levels to produce new interest rate scenarios that can be run through company models. Adding 3% to the current level of interest rates or subtracting 30% from the current level of equities is a deterministic approach that can provide insight into how sensitive a company may be to certain events. This process, however, lacks any insight into the probability of these shocks and therefore may leave practitioners with little guidance as to when or whether they should take risk mitigation actions.

Stress testing that includes probability distributions developed through the use of an ESG provides more actionable results. Using a set of scenarios produced by an ESG makes it possible to identify certain percentile changes and take appropriate mitigating measures that are commensurate with the likelihood of the event. A costly hedge to protect against a 1-in-1,000 event may be much less economical than the same hedge protecting against a more likely 1-in-20 event. Attention should be given to secondary factors in any particular scenario. There may be anomalies with the other factors that are specific to that one scenario, and any effects on the results of the model may mistakenly be attributed to the primary factor.

ESGs also allow for producing reverse stress tests, where the stress scenario is not selected before running the company model but is instead selected after producing the company results. Predetermining the scenario stress may prove not to be a meaningful stress test at all. More information may be gleaned from focusing on the scenarios that actually produce the poor company results. What is most impairing to an insurer may not be simply the worst interest rate shock, but instead a modest rate increase accompanied by a significant equity market drop and credit spread widening. Reverse stress testing helps identify problem scenarios and illuminates the most important risk drivers.

4.1.4 ECONOMIC CAPITAL

ESGs are an important tool for economic capital calculations. Economic capital focuses on the extreme tail events that can adversely impact the value or solvency of an insurer. ESGs provide a tool for quantifying the probability and scope of extreme events that may result from external economic and market volatility. Both market-consistent and real-world representations are important in developing economic capital calculations.

Economic capital is defined by a downside risk metric that quantifies adverse impacts on a company's capital, at a given probability level, over a specific time horizon. These adverse impacts are derived using company models that consider economic scenarios that produce extreme value changes. As with stress testing, a deeper appreciation of the risk factors that materially impact a company's financials is developed through understanding the combination of economic events such as inflation rates, unemployment and other factors, along with capital market events that affect interest rates and equity markets.

The specifications of the interrelationships of these risk factors are particularly important as companies seek ways of diversifying their risks for improved performance during adverse conditions. If the correlations among modeled factors are not reasonably represented, the diversification attributes will not be properly reflected.

Certain economic capital definitions require the calculation of market values at a future point in time, T_x . The scenarios used for traveling to T_x may be real-world representations that are not market consistent. However, within this path, the scenarios used to determine the market values at T_x may be market consistent.

Economic capital analysis focuses on extreme tail events that are often the result of broad economic crisis. Fixed-income investments may be considered to have low correlations with equities, but during periods of crisis, certain corporate bonds may become highly correlated with a falling equity market as spreads between corporate yields and government yields widen in sympathy with credit default concerns. Alternatively, government bond prices may spike during crises due to an investor flight to quality. When considering the tail events implied by these periods of crisis, particular consideration should be given to the number of scenarios used. If the concern is with the 99.5 percentile worst result, testing 200 scenarios would not provide sufficient confidence that the result is representative of the true risk. As the focus extends further into the extreme regions, even more scenarios would be needed. The number of scenarios needed to achieve a stable result will depend on the volatility of the results in the tail regions, and the use of 100,000 scenarios or more is not uncommon.

4.1.5 STRATEGIC ASSET ALLOCATION

ESGs using real-world scenario representations can support analysis for more efficient risk-adjusted investment allocations that consider the effect of liability correlations against a broad range of economic and financial market changes. The strategic asset allocation (SAA) process has evolved to a point where it relies heavily on the use of stochastic scenarios to test and validate new investment allocations relative to liability characteristics. The objective of this exercise is to find the most efficient allocation of invested assets for a given return objective and risk tolerance. However, economic scenarios that affect asset values may also affect liability values. Because of this, asset-only analysis will often produce efficient investment allocations that are significantly different from results that would be generated when liabilities are included.

An SAA may require the full capabilities of an ESG to produce a robust set of not only projected interest rate changes, but also projections of credit defaults and spreads, bond rating migrations, equity returns, inflation and other factors that may affect the way assets and liabilities behave. An SAA process evaluates the relative value of potential asset classes for investment along with the volatility and correlation of performance across all classes relative to the behavior of the liabilities.

Within asset classes, the performance of short-duration versus long-duration fixed-income securities is highly dependent on the yield curve's movement and shape, so the results of an SAA will depend on the robustness of the modeling of this projected movement, which would include the frequency of yield curve inversions and the volatility and skew representative of future movements.

Across asset classes, fixed-income credit spread movements should be related to, and consistent with, other assets classes such as equity performance. The behavior of credit spread changes should also represent realistic volatility and correlations, particularly in the extreme tail events, as any expected diversification value may be muted just when a company is relying on it the most.

The SAA economic scenarios should be applied consistently to assets and liabilities, regardless of whether different modeling software is used for evaluating these components. SAA analysis is performed using real-world scenarios as cash flow pricing is not the primary consideration. Mean reversion targets for the real-world scenarios need to be explicitly determined, as mean reversion dynamics are not implied as with market-consistent scenarios.

Mean reversion within an ESG model controls the movement of a variable to its assumed long-term steady-state level. The range of the simulations will then vary around the reversionary path from initial levels to the assumed steady-state level. There are several options for mean reversion targeting, such as reversions to long-term historical norms and remaining at current levels. Reversion levels could also target current long tenor rates with reversions toward historically normal yield curve shapes. For example, should the current yield curve be inverted or excessively steep, the yield curve projection may pivot at the 10-year yield to a longer-term normal upward-sloping shape. Equity returns may not revert to long-term historic levels, but the historic equity risk premium may be maintained, which would produce low returns relative to history but consistent returns relative to other asset classes in a low-interest-rate environment. SAA analysis often is designed to minimize the bias for or against any particular asset class. However, certain tactical insights can be gained by testing new reversion targets to reflect particular investor viewpoints.

4.2 PENSION AND RETIREMENT APPLICATIONS

In addition to similar applications previously discussed in the life insurance section, several specific applications in pensions are best understood through the use of an ESG.

Considerations regarding the number of scenarios, choice of market-consistent versus real-world scenarios, and specific calibration approaches may depend on the specific application being considered.

Pension industry applications include the following:

- Pension funding sensitivities
- Strategic asset allocation and liability-driven investments
- Pension risk transfer considerations

There are three major kinds of pension or retirement plans:

- Defined benefit (DB) plans, which guarantee a specific formula for retirement income, based on years of service and income record.
- Defined contribution (DC) plans, in which the sponsor and/or employee contribute a set amount of money to accumulate. At retirement, this accumulated value is transformed into retirement income but is not guaranteed in amount.
- Hybrid plans, which include a mixture of DC and DB plan features, often with a certain small portion of the retirement income stream guaranteed.

With respect to applications of ESGs, this section will primarily deal with DB plans and by extension will address some aspects of hybrid plans. It should be noted, however, that DC plans also can make use of ESGs specifically with respect to evaluation and design of target date funds. ESGs can be great tools for evaluating whether those designs can produce the required retirement income.

4.2.1 PENSION-FUNDING SENSITIVITIES FOR DB PENSIONS

In the case of a DB plan or even a hybrid plan, the funding amounts tend to be level percentages of payroll costs. In that case, the concern is that funding may be insufficient over time to support the benefits that have been promised. An ESG is essential for producing scenarios under which varying funding requirements are projected, especially when there are several interacting factors on both the asset and liability sides. While many actuarial pension-funding methods for DB pensions are geared toward having a predictable funding pattern, this works only if all pension valuation assumptions regarding payroll, investment performance and retirement elections or payouts are fulfilled. Sometimes actual experience of a pension plan varies from the valuation assumption, and these gains or losses must be made up over time with variances in funding.

Considering appropriate ESG model construction for pension analysis requires a full range of scenarios to understand performance in more extreme conditions, and a sufficient number of scenarios are required to produce and understand the likelihood of different levels of performance and to provide a basis for response. In projecting funding

requirements, pension sponsors are looking to see how supportable future funding requirements may be. Real-world scenarios are needed to determine likely paths that the sponsor may have to navigate.

The appropriate forecast horizon for pension scenario projections is a major consideration in pension ESG applications. The liabilities for life and annuity insurers can be very long, but pension liabilities can be even longer. Many life insurers may project their liabilities for a few decades in performing SAA or shorter time periods for economic capital modeling, but pensions last through working years and until either payout (or purchase of an annuity) or the death of the pensioner (and/or beneficiary, such as a spouse). One of the key aspects of pension-funding management is amortization of losses or gains from experience that varies from expected performance over a certain period of time. Many public pensions amortize over very long periods of time, extending beyond the retirement of the people being covered. Thus, an ESG that can project scenarios for decades may be required for pension fund modeling.

Even though pension payments can be paid out over several decades, plan sponsors are also focused on short-term financial implications. This requires balancing the need to model these long dated obligations with short-term issues like funding requirements.

The kinds of economic factors modeled for assets backing life and annuity liabilities also hold for pensions, but many pension funds have a greater diversity of asset classes than is typical for life insurance companies, because of less regulation limiting investment choices in the pension environment. On the liability side, broader economic factors such as wage inflation, consumer price inflation (for annuities with cost-of-living adjustments in retirement), and economic factors tied to employment, retirement and disability rates (if the pensions have disability-related options) should be considered. These will likely be correlated with factors affecting asset cash flows.

Finally, the financial health of the sponsor may need to be modeled. Whether the pension sponsor is a private company or a governmental body, an important consideration is whether it can make future funding requirements, or how the funding requirement will compare with modeled revenues. The sponsor's financial health, as well as the funding requirements of the plan, will need to be on a consistent basis for comparison.

To give a concrete example, many public pension plans found their asset values dropped as a result of the 2008 market crisis. Some pension plans that had been very well funded found themselves to be greatly underfunded—at the same time as tax revenues dropped in the ensuing recession. Many public plan sponsors could not make full pension contributions, because the sponsors were under great financial strain at the same time as their funds were. Some public plan sponsors issued pension obligation bonds (POB) to try to bolster funding, but many stressed sponsors found their own credit ratings affected, so their borrowing costs increased, diminishing the benefit of POBs. For private sponsors, financial health may be correlated with ESG-modeled economic factors such as broad market performance (if it is a publicly traded company) or specific sectors.

4.2.2 SAA AND LIABILITY-DRIVEN INVESTMENT (LDI)

ESGs need to consider a broad range of asset classes in developing effective asset allocation strategies for pension plans. Investment strategy is vitally important to pension plan sponsors. Most pension plans use ESGs when evaluating different investment strategies. Each alternative strategy can be evaluated by examining the distribution of outcomes for key pension metrics (contributions, expense, funded status, value at risk on funded status, etc.). The investment mix that generates most favorable outcomes to the sponsor based on the sponsor's objectives and risk tolerance becomes the optimal strategy.

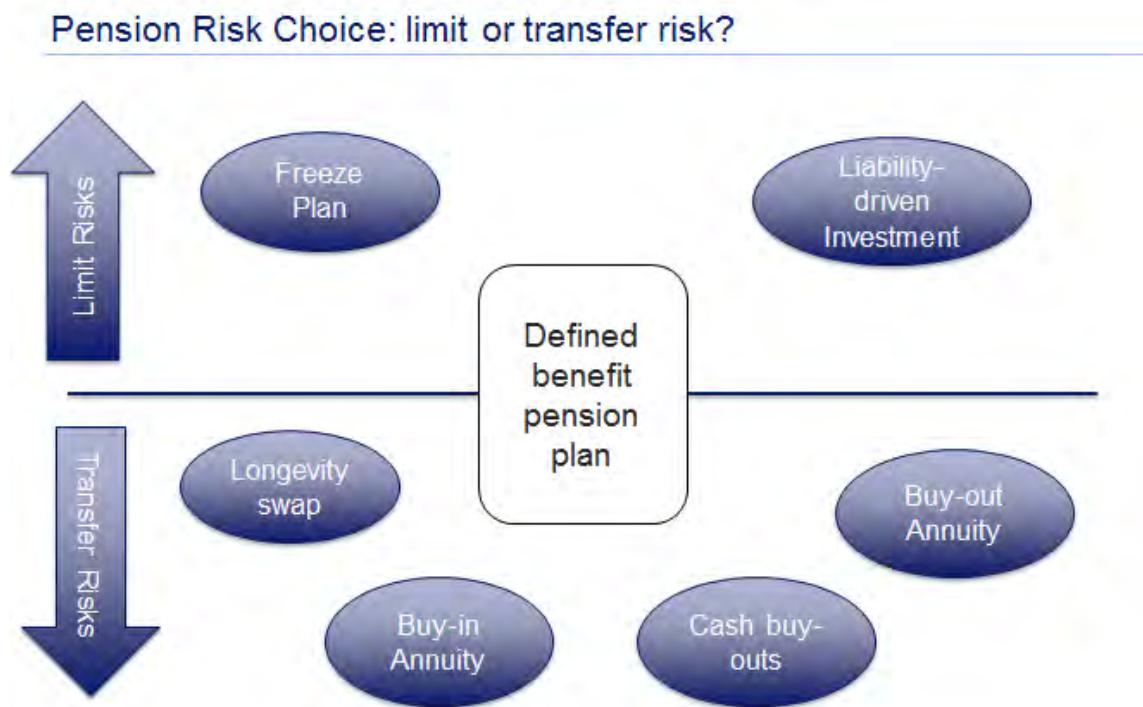
Liability-driven investment (LDI) is a particular type of investment strategy that defined benefit pension plans use in order to better match their assets with their liabilities. It has become a popular technique over the past decade. The main difference between pension-related LDI and insurer-related SAA is that the objective tends to be stability of contributions and expenses required under FASB and GASB guidelines, as opposed to maximizing enterprise value. Pension plans have traditionally held a great deal more in equities and other volatile asset classes than do insurance companies, because they are not as constrained as to what kinds of assets they can hold. One particular twist, however, is that some private pension plans hold stock of the sponsoring company. There are regulatory limits to this, but one important aspect to note is that the performance of any such stock held will be completely correlated with the sponsor's financial health (and thus ability to fund).

For open plans, investment strategy tends to favor growth assets like stocks as a way to provide for ongoing benefit accruals. In contrast, plans that are closed to new entrants and benefit accruals are generally looking to eventually terminate and are therefore more concerned with better matching their assets to their liabilities and minimizing the chance of asset/liability shortfalls.

4.2.3 ESGS SUPPORTING PENSION RISK TRANSFER CONSIDERATIONS

Many DB pension sponsors have come under stress in the past decade, partly due to asset-side underperformance and partly due to unexpected liability-side increases, such as increased longevity. As a result, sponsors have been looking at their options to reduce their economic exposure to the variability of results of their pension plans. Some of the choices can be seen in Figure 4.1.

Figure 4.1



ESGs can be useful to sponsors looking to compare their options, especially with options where they are still completely in control of the pension plan: choosing to freeze the plan, implementing LDI (whether plan is open or frozen) against options that involve transferring the risks to third parties.

Some of the risk transfer options, such as cash buyouts (lump sum settlements) and buyout annuities, completely remove the pension obligation from the sponsor's balance sheet. Cash buyouts are subject to minimum interest rate and mortality assumptions to calculate lump sum cash-out amounts in conformance with government regulations. Buyout annuities are priced by insurers with risk premiums reflecting the size and nature of the risk being transferred.

The other two options—longevity swaps and buy-in annuities—leave the pension obligation on the sponsor's balance sheet but provide assets that are supposed to be somewhat correlated to the pension liability. In the case of longevity swaps, specific non-economic risks may be removed, but the risks from economic factors remain, so using ESGs to model future funding requirements would be important in evaluating this choice.

4.3 PROPERTY AND CASUALTY (P&C) APPLICATIONS

Applications of ESGs in property and casualty insurance are often focused on the impact of inflation on liabilities and assets, and on economic cyclical characteristics affecting both exposures and policy pricing.

Property and casualty modeling tools and methodologies have evolved along somewhat different paths compared with life, driven by differences in the fundamental risk factors and time horizons involved. There is, however, significant common ground between P&C and life insurance companies in terms of the need for economic scenarios in risk analysis. Foremost is the fact that both types of companies carry substantial portfolios of invested assets, which are exposed to variability in the financial markets due to movements in interest rates, credit spreads, equity prices, etc., as well as the risk of outright default. The kinds of activities for which real-world scenarios would be desired would be the same in both life and P&C, and similarly for market-consistent models.

However, there are material differences that have implications for the design and calibration of an ESG. Several specific applications in property and casualty insurance are best understood through the use of an ESG.

Considerations regarding the number of scenarios, choice of market-consistent versus real-world scenarios, and specific calibration approaches may depend on the specific application being considered.

Property and casualty products have different characteristics with respect to liability development, with many casualty products involving significant time lags in settlement or even discovery. Property products are often volatile based on factors such as weather events, crime trends and economic supply-and-demand characteristics that may require greater liquidity reserves on the asset side. Because of the complexity of the interaction of these factors over an extended time horizon, an ESG provides a much more comprehensive tool to understand the range of potential outcomes. Today an ESG is considered a necessary component of enterprise risk modeling for a P&C insurer.

ESG-driven applications in property and casualty insurance include the following:

- Liability valuation
- Economic capital and stress testing
- Pricing cycle and policy management
- Strategic asset allocation and liquidity

4.3.1 LIABILITY VALUATION ISSUES

Liability valuation issues are related to inflation and the time horizon.

Inflation. P&C insurance claim costs are generally sensitive to price inflation, either in the economy overall or in specific subsectors of the economy such as health care, the housing market or construction. P&C actuaries provide for the expected rate of claim cost inflation in the rate estimation process, but there can be significant uncertainty in these estimates. Even for products with short claim payment tails, there can be over two years' difference between the historical experience period that is used as a basis for the rates and the future time at which the claims will actually be paid. For longer-tailed products, this difference can be five, 10 or more years.

Further, there is often a delay in recognition of adverse inflationary impacts even after they have occurred, which means that years can pass before rates will be corrected for those impacts. Thus, several years of underpriced business can accrue before the financial damage can be mitigated. Similar to inflation, risk arises due to the time lag between the time premiums are estimated and the time the insurance claims are paid. For example, in some occurrence-based liability lines, the benefits (claims payment) may be subject to cumulative inflationary effects and may lag the time that rates are estimated by years or even decades. Pricing and other underwriting decisions may exhibit considerable volatility as a result of needing to respond to these lagging signals.

The implication for an ESG is that the models of future inflation scenarios must be robust in their ability to create tail scenarios with very high inflation rates and with a realistic mix of inflationary patterns over time.

Time horizon. Life insurance projection models typically need to cover future time horizons of several decades, due to the nature of the benefits provided. While some P&C insurance products (notably workers' compensation) can have comparably long liability horizons, for most products, the bulk of the claim payments occur within a few years of the policy period. Loss reserves may be set and reset on an annual basis or even more frequently. Therefore, in many cases, the calibration of an ESG for a P&C insurer can focus more on stability and robustness in the first few years of the projection period, perhaps at the expense of longer time horizons.

4.3.2 ECONOMIC AND CAPITAL ANALYSIS AND STRESS TESTING

An ESG is essential for performing a robust analysis of economic capital requirements for property and casualty companies. All of the comments in the earlier section on economic capital analysis for life insurers (4.1.4) apply to P&C insurers as well. As we stated in that section, economic capital considerations generally refer to tail events, usually in terms of an economic or financial markets crisis. For economic capital analysis, at least in the context of solvency assessment, risk metrics typically focus heavily on relatively extreme downside percentiles of the distribution of outcomes.

Risk factors and stress considerations differ from life insurance considerations. Inflationary factors can compound over time, and while they may correlate with interest rate discount factors that would affect both asset and liability cash streams, loss cost inflation may differ materially from interest rates in magnitude in many scenarios.

While financial stresses are the most salient factors stressed in life-annuity stress testing, some P&C insurers are subject to the risk of major natural and human-made catastrophic loss events, such as hurricanes, earthquakes or terrorist attacks. In addition to implications on the need for liquidity to cover such events, they could have follow-on effects on prices in the financial markets, and on inflation in areas that affect P&C insurance claim costs (notably construction costs).

Studies have found that such effects have been very minor and transient, at least in big economies; however, they could be more material in smaller economies or in the event of a catastrophe worse than those seen in the past.

While there is debate as to the likelihood of such effects occurring, the possibility should be contemplated at least in the definition of stress tests.

A more important concern related to catastrophes (and possible reinsurance credit issues) relates to liquidity issues. Liquidity concerns may be a function of the potential for extreme liquidity correlations across financial markets, capital markets and reinsurance markets in the event of extreme catastrophes.

4.3.3 PRICING CYCLE AND POLICY MANAGEMENT ISSUES

ESG models can be used to inform product pricing by helping to establish the profit margins needed for each product to compensate the insurer for the need to carry capital on the balance sheet to support them. ESG models can support the complexity of the demands on risk capital from multiple product lines and business lines in recognizing opportunities that these areas produce to diversify risk.

Policy time horizons are generally annual, as opposed to the longer-term policies that are typical in life products. Repricing of exposures is usually performed annually, though loss reserves may build up over a long period before they are paid. Loss reserves may be set and reset on an annual basis or even more frequently. Management may respond to significantly changing conditions (including economic conditions) with pricing, loss reserving or other adjustments that may change the medium-term outlook for the business, and some model applications may contemplate these possibilities.

The ESG is just as important in this area as in economic capital analysis, but there can be some distinctions, depending on how risk is defined for performance measurement purposes. For economic capital analysis, at least in the context of solvency assessment, risk metrics typically focus heavily on relatively extreme downside percentiles of the distribution of outcomes.

However, for performance measurement purposes, a company's management may also be concerned with other parts of the distribution. For example, risk may be measured in terms of the overall volatility of results as measured by standard deviation, or based on percentiles closer to the center of the distribution.

While there can be variations in how these metrics are constructed, a common characteristic is that business segments with greater risk will need to have a higher level of expected profitability in order to be financially attractive. This may have implications for how an ESG is calibrated and validated.

4.3.4 STRATEGIC ASSET ALLOCATION INCLUDING LIQUIDITY CONSIDERATIONS

As P&C liabilities are not directly linked to invested assets, unlike with life-annuity products that have financial guarantees, and P&C insurer balance sheets are less leveraged in general, the liabilities held by P&C insurers are a diversifier, as opposed to amplifier, of risk for the insurers. This enables P&C insurers to take on riskier assets without an unacceptable risk to capital. An ESG application may provide for a more efficient asset allocation approach that can increase income or lower overall risk capital requirements.

That said, due to catastrophic risk exposure, P&C insurers can have a liability volatility of such high magnitude that this creates potential liquidity events. In an SAA exercise, the modeler may have to impose certain liquidity requirements or consider liquidity stress testing after an SAA portfolio is chosen.

On the asset side, modelers may consider including assets classes such as municipal bonds and inflation-linked bonds. The first provides a tax advantage for P&C insurers, and the second provides some hedging to the inflation-linked risk of the liabilities.

However, the credit risk dynamics for municipal bonds are notably different from those for the corporate credit markets. General obligation bonds are usually backed by the full faith and credit of the municipal entity, and the funding for some come from dedicated taxes that may be less affected by the credit cycle. In recent years, there have been some high-profile municipal bankruptcies, even though the corporate bond default rate has been very low. As a result, specialized credit models can be important in properly quantifying the risk exposure of a P&C company with material amounts of these bonds in its portfolio.

4.4 CHAPTER SUMMARY

1. ESGs are important in a wide variety of business decision-making applications within the financial services industry beyond those applications that are required for regulatory compliance.
2. Key application-specific considerations must be addressed when using ESGs. These include the use of real-world or market-consistent scenarios, steady-state behavior, the number of scenarios (or scenario subset selection) and the model projection period and time horizon.
3. ESG model requirements differ for different business uses across life, P&C and pension businesses.
4. Comprehensive ESG modeling capabilities are often required; these include credit, inflation and other economic and capital market variables, in addition to interest rates.

Chapter 5: What Are the Essential Features of a Good ESG?

This chapter discusses the essential features of economic scenario generators and then distinguishes between real-world and risk-neutral applications. Several examples are used to illustrate the depth and rigor required.

As was described in Chapter 1, an economic scenario generator (ESG) is a software tool that generates a collection of simulated economic scenarios that represents a distribution of possible economic futures. The role of an ESG is to provide forward projections of economic and capital market variables in support of risk management and pricing activities. Real-world scenarios are concerned with forward-looking potential paths of economic variables and their potential influences. Market-consistent, or risk-neutral, scenarios are concerned with mathematical relationships within and among financial instruments.

The actual design, operation and maintenance of an ESG are highly technical topics. The specialized topics and associated criteria essential for having a good ESG are not easily distilled to a few sentences. Nonetheless, it is generally expected that an ESG should meet the following minimum requirements:

- Production of simulation results that reflect a relevant view
- Inclusion of some extreme but plausible results
- Scenario generation that embeds realistic market dynamics

As this chapter will illustrate, obtaining sufficient dispersion by way of a collection of future economic possibilities is an essential requirement for capturing market risk. However, an ESG is not a predictive tool used to forecast future values of economic variables. It is thus important to distinguish an ESG from (a) econometric models that specify direct algebraic relationships between various economic quantities and that are intended for statistical analysis and forecasting, and (b) general equilibrium models that attempt to predict the effect of changes in key economic variables and are often used in economic policy analysis.

Section 5.1 of this chapter gives an overview of ESG applications with concentration on real world versus market consistent. Real-world simulation captures market dynamics and risks, and applications would include simulating interest rates and testing portfolio investment strategies. Market-consistent simulation focuses on valuing cash flows, and applications would include pricing of guarantees or derivatives.

Section 5.2 represents the bulk of this chapter. It outlines the essential features of a good ESG model. A few of the key features outlined in this section include reflecting a sufficient range of financial markets and economic factors to meet the purpose for which the simulations are being applied (including regulatory requirements). An ESG should balance practical limitations of being computationally efficient and numerically stable, while still reflecting the possibility of extreme events. It should provide sufficient output detail for extensive validation and accommodate many types of calibration views. And it should have an underlying structure that maintains sound and consistent logic across models (including both real-world and market-consistent simulations) based on an understanding of the way financial markets realistically interrelate.

Section 5.3 briefly deals with advanced ESG model attributes by way of examples of how an ESG needs to reflect some more extreme aspects of the financial markets. These would involve advanced technical considerations to capture more specific market risks that may be important for some applications. Two examples of advanced attributes are capturing corporate bond spreads that reflect a liquidity component even in high-grade corporate bonds, and allowance for negative nominal yields in Treasury yield dynamics.

Section 5.4 ends with the question of how good is any ESG, really. It addresses the question of whether extreme but plausible events, which a good ESG should cover, can be expected to cover all possible events. A good ESG should permit sufficient recalibration capabilities to extend simulation to alternate sets of extreme, but still plausible, events.

5.1 REAL WORLD VERSUS MARKET CONSISTENT

The application of an ESG can range from smaller tasks, such as the simulation of interest rates in order to understand the impact of interest rate changes on an investment portfolio, to a broader role such as testing a range of portfolio

investment strategies across a wide range of asset classes. ESG technology is also used in the risk management of investment guarantees such as those that are commonly bundled with life insurance products. Such a risk management application might involve both pricing the guarantees and measuring the effectiveness of a hedging program for the contingent obligations associated with the guarantees.

The scope of an ESG application thus can vary widely. The examples in the preceding paragraph include both real-world and market-consistent applications. The simulation of interest rates and the portfolio investment strategy testing are examples of real-world simulation. The pricing of guarantees is an example of market-consistent simulation. The measurement of the effectiveness of a hedging program will often involve the application of both real-world and market-consistent simulations.

Real-world simulation is a concept that can be intuitively understood. The purpose of a real-world simulation is to capture market dynamics, risks and returns in a way that an insurance company or other financial institution might experience them. Real-world simulations enable the exploration of the what-if questions asked by management as it tries to gauge the likelihood of future events and their business impact.

Market-consistent simulation is less intuitive and has been a source of confusion from the very beginning of the application of financial risk management. Market-consistent scenarios provide a tool for valuing cash flows. Generally, the situations where market-consistent scenarios are needed are those in which one is pricing cash flows that depend upon stochastic financial variables. A classic example is the need to determine the economic cost of an investment guarantee such as a guaranteed minimum rate of return on an equity or fixed-income portfolio.

As this chapter will discuss in more detail, an ESG should be able to accommodate a broad range of applications involving both real-world and market-consistent scenarios. Users of an ESG will need to have a good understanding of the specific applications they need to address and to be able to determine whether these applications require real-world scenarios, market-consistent scenarios or both.

5.2 ESSENTIAL FEATURES OF A GOOD ESG

There are many essential features of an ESG. In the list that follows, we identify the main features that we believe make the difference between a basic ESG and one that is truly powerful in its ability to uncover possible risk scenarios.

To clarify some terminology in advance, the discussion that follows references parameterization and calibration. In the context of an ESG, the parameters are the part of the ESG that can be changed and that governs the dynamics of an economy. A particular choice for the parameters is referred to as a parameterization. An ESG can have several hundred parameters—mean, standard deviation, correlations, mean reversion speed, "jump" behavior, to name just a few—that make it possible to produce future paths that have features and qualities consistent with the desired targets. Poor parameterization will result in a model that does not exhibit behavior consistent with the desired targets.

Calibration is the process of setting the parameters of the ESG. A calibration may be done with the primary objective of fitting historical targets, or it may be focused on expressing a particular view on the economy. The methods of calibration vary depending on the primary objective and, in practice, will include statistical techniques such as maximum-likelihood estimation and the fitting of the model to market data such as yield curves. In some instances, the user of an ESG may choose to calibrate a model in a manner that is not consistent with the behavior of the historical record. This could result from a user looking to voice a view on the economy and assess the risks associated with that view. A good ESG has sufficient parameters to control the key dynamics that the user wants to control, and a good ESG will have sufficiently rich dynamics to capture the aspects of the data that the user wishes to mimic.

A good ESG is thus a function of both its model dynamics and its parameterization. While a good set of simulated economic scenarios requires both a good ESG and an appropriate parameterization, a good ESG can have a bad parameterization, and a bad ESG can have a good parameterization. One will generally assess the quality of the simulated economic scenarios by validating them against the primary calibration targets.

5.2.1 A SOUND FOUNDATION, BALANCING PRACTICALITY AND COMPLETENESS

A good ESG has a sound foundation for the way the models are built and the way the variables are interrelated; it balances practicality and completeness. One might consider an ESG to be a collection of mathematical models that fit the data. However, a good ESG will construct these models with some economic logic based on an understanding of the way key features of the financial markets relate to one another. For example, a good ESG will capture leverage effects in equity markets, which might require modeling how the dynamics of volatility affects drawdown risk. In contrast, a highly overparameterized vector autoregressive (VAR) model¹ might fit the data but lack any well-thought-out dynamics that can be trusted to produce the extreme but plausible scenarios that are so critical in risk management.

A good ESG strikes a balance between practicality and the ability to capture the critical features of the historical record. It is not practical to model every variable and every outcome precisely, so one must determine which attributes are the most important to capture. Models that are too simple, however, can miss the true risks that affect a business. Simple models do not capture the likelihood and extent of extreme events that could result in expensive or ruinous mistakes. At the other extreme, a good ESG does not “overfit” the data.² Overfitting reduces a model’s ability to explain or robustly simulate market behavior.

5.2.2 A SUITE OF MODELS SUFFICIENT TO CAPTURE IMPORTANT ASSET CLASSES AND ECONOMIC VARIABLES

A good ESG provides a suite of models that is sufficient to capture the asset classes and economic variables of greatest importance to the risk profile of a given firm. Most ESG applications will make use of the following components for each economy where there is exposure:

- Sovereign interest rates and returns
- Equity returns
- Inflation
- GDP

Some economies will include additional components such as these:

- Unemployment
- Corporate bond yields and returns
- Mortgage-backed securities
- Covered bonds (German Jumbo Pfandbriefe)
- Municipal bonds (important in U.S. models)
- Derivatives

A multi-economy ESG simulates all of the single-economy variables and will also take into account realistic interactions between economies. Among many others, these interactions include considerations of foreign exchange, correlations between the stock indices of two economies, and dynamics of correlations that tend to increase during bear markets.

A good ESG also provides additional modeling capabilities so that the available asset classes and economic variables can be expanded in a consistent manner to meet more specialized user requirements. For example, a user may require a total return index for an asset class that is not standardly included in the ESG; additional modeling capabilities permit the user to construct such a total return index using available ESG variables.

¹ VAR models are econometric tools that are used extensively in economics. A strong case for the use of VAR model in econometrics was made in Sims (1980) on the grounds that they do not require a detailed theory to estimate relationships between economic variables. Campbell and Viceira (2002) provide many examples of the application of VAR techniques in an investing context.

² When models are built, whether an ESG model or not, a key consideration is that the model should be able to accommodate important features of the historical data used to estimate the model. Another critical consideration is that the model should be able to produce realistic simulated output or forecasts beyond the historical data. If one forces the model to tightly fit the historical data, then the robustness of the model output outside of the historical data is likely to be restricted. This is the notion of overfitting. When a lot of parameters or intricate relationships are imposed on a model to achieve a tight fit to the historical data, then the model is at risk of becoming overfitted. Leinweber (2007) and Bailey et al. (2014) provide examples of overfitting in financial contexts.

5.2.3 ABILITY TO ACCOMMODATE MANY TYPES OF CALIBRATION VIEWS ACROSS A WIDE RANGE OF BENCHMARKS

A good ESG is capable of accommodating many types of calibration views across a wide range of benchmarks. There are many different views of the global economy that an ESG might be called on to reflect. For instance, an ESG might be calibrated to a view that embeds historical benchmarks for the past 50 years. Such a view would accept the statistical features of a 50-year data window as the calibration benchmark.

Another calibration view might consider short-term and long-term targets, and these targets would recognize historical norms but would be set with some measure of expert judgment. Such a calibration view typically would consist of statements such as this:

One year into the simulation, the average 10-year Treasury yield should be 2.6%, and the long-run average 10-year Treasury yield should be 4.75%.

Another example might involve a calibration to reflect historical dynamics but to adjust mean expectations to be consistent with current consensus opinions. Other calibration views might embed externally mandated calibration views that require minimum tail behavior on equity or fixed-income returns. This sort of calibration view is routinely required by the American Academy of Actuaries, the Canadian Institute of Actuaries, and regulatory authorities such as the Office of the Superintendent of Financial Institutions (OSFI) Canada.

Yet another example of a calibration view would be to impose a range of interest rate behavior over the near term. For example, one might require that interest rates remain at low levels for the first two years of the simulation before gradually reverting to a higher longer-term level. In such a calibration view, one stops short of defining a known interest rate scenario but seeks to impose a strong requirement in order to test a specific business objective.

5.2.4 SIMULATION RESULTS THAT REFLECT A RELEVANT VIEW

While being capable of accommodating many types of calibration views, a good ESG produces simulation results that reflect a relevant view. When calibrated to reflect a certain view, a good ESG will produce simulation results that are relevant to historical facts. A common tendency is to overweight the nature of the recent past. Indeed, there is an important difference between, on the one hand, setting a view and then permitting realistic model dynamics to take over and, on the other, setting a view that overwhelms any realistic market dynamic. One must recognize that the headline problems of the day have a strong tendency to dominate the modeling and thus the calibration process. This is understandable because users of an ESG are intensely focused on their business risks, and these are generally aligned with the headline problems of the day. The danger in placing too much focus on headline risk is that one can forget that, over moderate time periods, the economy can migrate to new and different worrisome places. For a risk manager and risk management application with a longer-term horizon, one must avoid the temptation to over-impose a view.

5.2.5 SOME EXTREME BUT PLAUSIBLE OUTCOMES

A good ESG produces some extreme but plausible outcomes. A careful study of historical data teaches us that the financial markets have experienced a great many interesting and extreme events. The notion that a good ESG should produce some extreme but plausible scenarios is intuitive and reasonable, but it is a broad concept rather than a blueprint for building an ESG. By “extreme scenario,” we mean a scenario that reaches or surpasses the limits of experience. By “plausible scenario,” we mean a scenario that conforms to economic principles and can be rationalized in an economic context. A scenario for which long-term government bonds return more than equities or corporate bonds is plausible. A scenario for which inflation averages 20% and long-term Treasury yields average 10% is not plausible.

The inflation and interest rate examples in the previous section illustrate the requirement for an ESG to produce some extreme but plausible scenarios. Recent experience with protracted periods of low interest rates has now established that the short end of the Treasury yield curve can remain low for extended periods of time. This has been evident in Japan for many years and has also been a feature of the U.S. and European markets for the past several years. A good ESG should be able to produce scenarios that exhibit this feature while still allowing for scenarios with higher interest rates.

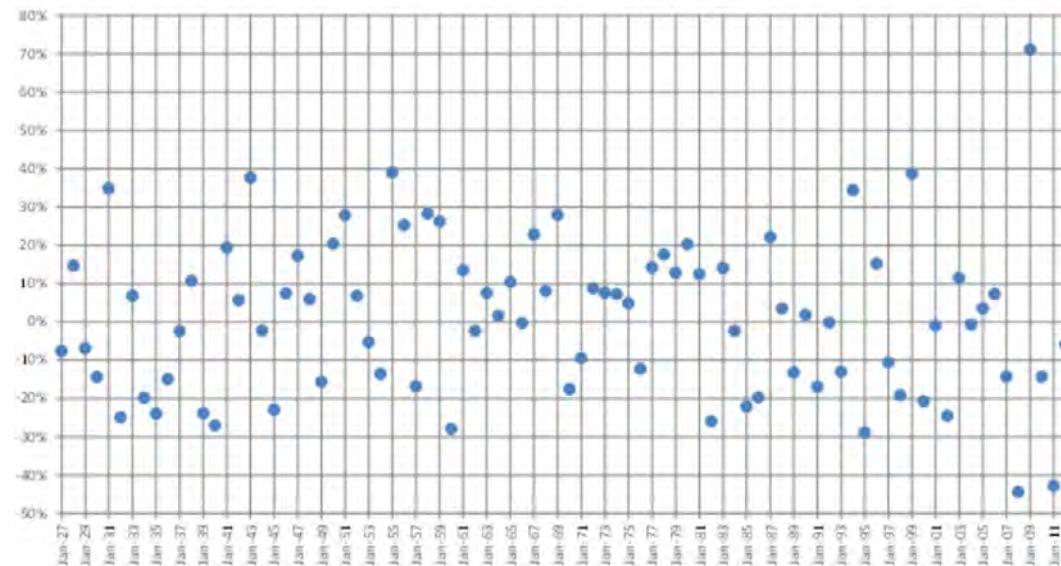
One measure of the ability of an ESG to meet or exceed historical norms is to compare simulated returns with historical returns and observe whether the simulated returns have heavier tails than history. A good ESG generally will have this property. However, other concepts of extreme outcomes should be considered.

For example, a good ESG not only produces realistic distributions for key economic variables but it also represents the correct “pathwise” behavior, that is, how the simulation moves through time to produce a particular outcome. The pathwise behavior reflects the course of the variable in periods of higher or lower volatility as well as “jump” behavior (large changes in price movements from time to time) and volatility clustering (tendency for large or small changes in prices to cluster together).

Returning to the topic of interest rates, Figure 5.1 shows the annual percentage change in U.S. 10-year Treasury yields for the period 1926 through 2012. Until the financial crisis, these relative changes tended to range between –30% and 40%. The behavior of 2008, 2009, and 2011 represents extreme outcomes relative to the historical record. A good ESG is capable of producing annual percentage changes in Treasury rates that capture historical norms. It should also be capable of producing drawdowns and run-ups that exceed these same historical norms.

Figure 5.1

ANNUAL PERCENTAGE CHANGE, U.S. 10-YEAR TREASURY YIELDS, 1926–2012



Sources: Ibbotson; Bloomberg; Conning, Inc.

5.2.6 REALISTIC MARKET DYNAMICS EMBEDDED IN THE MODEL

A good ESG embeds realistic market dynamics. Market dynamics are not merely statistical targets but attributes that reflect volatility or the way a variable changes and relates to other financial variables. The idea that a good ESG should embed realistic market dynamics is a guiding principle rather than a specific modeling feature. Indeed, market dynamics vary greatly across time, and no model can capture them all. When one talks about market dynamics, one is really asking about how markets work. A vast amount of research is done in this vein every year, and the conclusions often change over time. However, experience has shown that there are certain market dynamics that figure prominently in risk management and are essential for a good ESG. These include yield curve movements, changing relationships between asset class returns, return correlations, corporate bond migration and default, and corporate bond spread dynamics. As illustrative examples of market dynamics, consider the following:

- Yield curves can shift and twist in many directions, and a wide range of movements is possible.
- Over the longer term, there will be periods when one asset class outperforms, followed by periods when that asset class lags.
- Over extended periods of time, the strength of the correlation of returns between asset classes changes.

- Corporate bonds may exhibit periods of fairly predictable migration patterns and then suddenly become unstable.
- Corporate bond spreads are subject to periods of extreme spread contraction and widening.

Before turning to the examples of market dynamics, we first address the concept of a “path.” A single simulated scenario consists of the outcomes for each economic variable, starting from its initial value and running from the first simulated value to the last simulated value. For instance, a simulated scenario will specify what happens to interest rates after one simulation period, two simulation periods, and at the last simulation period that occurs at the end of the simulation time horizon. We follow the standard practice of referring to such a single simulated scenario of all financial variables as a path. The significance of the simulated path is that it represents one possible future evolution of the economy and therefore represents one possible future complete economic experience.

Taken together, a collection of simulated paths represents the simulation output from the ESG. When one looks at statistics such as average interest rates five years into the simulation, these are statistics that aggregate behavior across all paths. Equally important are measures that relate to the attributes associated with a specific path, with fluctuations in correlation and volatility being classic examples.

A good ESG should produce pathwise behavior that reflects market realities, because it is along a path that an insurance company experiences the evolution of the economy and responds with management decisions. Many return series have periods of both elevated and subdued volatility—a pathwise effect. If one considers only the distribution of returns, then one could be using a model that has an acceptable return distribution but fails to recognize the volatility characteristics that may or may not be reasonable. A good ESG should incorporate dynamics that produce pathwise fluctuations in realized volatility, because pathwise fluctuations in realized volatility affect investment allocations, hedging costs and mark-to-market positions.

5.2.7 COMPUTATIONAL EFFICIENCY AND NUMERICAL STABILITY

A good ESG is computationally efficient and numerically stable. An ESG is a complex system of models, underlying variables and parameters. Choice of models should take into consideration those with properties that make them efficient to implement computationally—i.e., faster run times and greater accuracy. Choice of models should also be robust enough to maintain stable parameter estimates.

A good ESG is calibrated over a broad enough range of history to reflect actual market dynamics. Parameter estimates for the models should exhibit some measure of robustness to the data window on which they were based. They should faithfully capture changing initial conditions while, at the same time, having underlying parameters that typically evolve at a moderate pace.

Although there is an ongoing need to refine parameter estimates in light of evolving market risks, one generally seeks to maintain a degree of stability in the parameter estimates. First, robust ESG models naturally tend to have stable parameter estimates as new data are added to a large historical data window. Second, making significant parameter changes from estimation period to estimation period is disruptive to the broad ERM process, because it would induce notable changes in the risk profile of various asset classes.

To be computationally efficient, a good ESG will take advantage of closed-form or semi-closed-form pricing algorithms that reasonably reproduce market prices. ESGs without such features are cumbersome to estimate from market data, slow to simulate and difficult to recalibrate to a user’s views. A good ESG must make full use of efficient computational algorithms for pricing to ensure both good run time performance and numerical accuracy.

5.2.8 CAPABILITIES AND CONSISTENT MODELS FOR REAL-WORLD AND MARKET-CONSISTENT SIMULATIONS

A good ESG has capabilities for real-world and market-consistent simulations and uses consistent models across both modes. Our discussion has primarily focused on the essential attributes for a real-world ESG. Comprehensive risk management, however, also requires the pricing capabilities of a market-consistent ESG. For example, if one is interested in real-world simulations that require derivatives prices or embedded value calculations, then the same suite of models should be used to maintain consistency, with appropriate parameter adjustments for the risk-neutral

and market-consistent modes. To apply “nested stochastics,”³ one applies a risk-neutral version of the ESG within a real-world simulation. For this nested stochastic process to make sense, the ESG must have real-world and risk-neutral versions available. Some ESG models are limited to real-world simulations, while others are limited to risk-neutral simulations. Consequently, if it is known that a nested stochastic process is required, then ESG models with real-world and risk-neutral versions must be available.

Some models are useful for trader-style pricing but do not have real-world dynamics. This is a significant limitation on holistic risk management, which is the goal of an ERM approach. Indeed, pricing models that do not have real-world dynamics might generate useful pricing, but one is then faced with the challenge of moving these prices into a real-world context to generate realistic risk assessment, and this must typically be done in an ad hoc fashion. Further, the use of a core set of models for both real-world and risk-neutral applications facilitates the greater satisfaction of regulatory requirements that management validate the use and effectiveness of internal models used for economic capital and ERM (the “use test”).⁴

5.2.9 COMPLIANCE WITH THE REQUIREMENTS OF REGULATORS AND AUDITING FIRMS

A good ESG meets the requirements of regulators and auditing firms. Given the heightened emphasis on risk and capital management within the insurance industry, a good ESG meets the growing demands of regulators and auditing firms. ESG methodologies should be transparent to support regulatory requirements. This entails documentation of the maintenance processes (e.g., parameter and software updates) and other supporting materials that enable a comprehensive understanding of the theoretical and empirical integrity of the simulated scenarios.

5.2.10 SUFFICIENT SIMULATION DETAIL FOR EXTENSIVE VALIDATION

A good ESG produces sufficient simulation detail for extensive validation. Strictly speaking, the process of validating an ESG is not part of an ESG. While it is self-evident that a good ESG validates to its calibration benchmarks, the process of performing the validation can require extensive inputs. A good ESG is capable of providing very detailed simulation output that is needed to validate its performance. While one expects a good ESG to provide a range of total return and yield output, one should also expect that a good ESG provides detailed output on default behavior, rating transitions, and the discount factors needed to perform regulatory risk-neutral testing.

5.3 ADVANCED ESG ATTRIBUTES

The previous section gave some relatively general characteristics that form a foundation for a sound ESG. An ESG is a complex system. If one is prepared to entertain a higher level of technical discussion, additional detail can be provided on other specific model attributes that are essential for capturing market risks. This greater depth focuses on very specific attributes of the asset classes being modeled. We mention two just briefly.

First, the corporate bond liquidity premium is an important practical modeling issue. Much has been written about the fact that a significant percentage of the spread on a high-grade corporate bond is due to something other than default risk. A good ESG is capable of capturing spreads that are wide enough to be realistic but that do not imply unrealistic default rates. This requirement means that a liquidity component should be included in high-grade corporate bond spreads.

Second, the finer aspects of Treasury yield dynamics include the allowance of negative nominal yields. It is now well established that negative nominal yields are possible, given both the recent experience in the U.S. and Germany and the negative nominal yields experienced in the United States during the 1930s and 1940s. A good ESG is able to produce negative nominal yields.

³ Nested stochastics or inner loops is a simulation approach where variables are projected using a real-world calibration (outer loop) and pricing is performed at each future real-world node, using a second stochastic simulation (inner loop) that is based on a market-consistent calibration. This simulation technique is used when pricing of contingent cash flows is required as part of a risk management application.

⁴ The “use test” is the requirement that an insurer demonstrate that its risk management models are an integral part of its management process. More detailed descriptions of the concept may be found in Solvency II documents such as CEIOPS’ Advice for Level 2 Implementing Measures on Solvency II: Articles 120 to 126 - Tests and Standards for Internal Model Approval.

5.4 HOW GOOD IS AN ESG, REALLY?

One of the nagging dilemmas faced by senior management relying on ESG technology is this fundamental question: How good is it, really? The preceding sections have laid out elements of the foundation for answering such a question. These include the ability of an ESG to perform the following tasks:

- Produce simulations that reflect a relevant calibration view
- Produce extreme but plausible results
- Generate scenarios that embed realistic market dynamics

However, what considerations might these foundations have overlooked? A relevant and practical testing ground for this question is the financial distress that markets periodically endure. While we have gone through periods of relative calm and relative panic since the financial crisis began with the meltdown of bank structured investment vehicles (SIVs) in the summer of 2007, it is reasonable to suppose that we have not witnessed the most extreme events that might be possible or even plausible. The stark reality is that no ESG can model all possible events, but a good ESG should produce a rich set of extreme but plausible events. A good ESG will also permit recalibration to different attributes in order to provide alternate sets of extreme but plausible events.

5.5 CHAPTER SUMMARY

This chapter has provided a high-level discussion of the essential features of a good economic scenario generator. An ESG is a complex system that is generally evolving in response to changes in market fundamentals and regulatory requirements. A good ESG has general characteristics that include the following:

1. It provides a sound foundation for the way the models are built and the way the variables are interrelated; it balances practicality and completeness.
2. It provides a suite of models sufficient to capture the asset classes and economic variables of greatest importance to the risk profile of a given firm.
3. It is capable of accommodating many types of calibration views across a wide range of benchmarks.
4. While being capable of accommodating many types of calibration views, it produces simulation results that reflect a relevant view.
5. It produces some extreme but plausible outcomes.
6. It embeds realistic market dynamics.
7. It is computationally efficient and numerically stable.
8. It has capabilities for real-world and market-consistent simulations and uses consistent models across both modes.
9. It meets the requirements of regulators and auditing firms.
10. It produces sufficient simulation detail for extensive validation.

5.6 CHAPTER REFERENCES

Bailey, David H., Jonathan M. Borwein, Marcos López de Prado, and Qiji Jim Zhu. 2014. Pseudo-Mathematics and Financial Charlatanism: The Effects of Backtest Overfitting on Out-of-Sample Performance, *Notices of the AMS* 61(5): 458–471.

Campbell, J., and L. Viceira. 2002. *Strategic Asset Allocation*. New York: Oxford University Press.

Leinweber, David J. 2007. Stupid Data Miner Tricks. *Journal of Investing* 16(1): 15–22.

Sims, C. 1980. Macroeconomics and Reality. *Econometrica* 48(1): 1–48.

Chapter 6: Considerations of Model Specification and Stylized Facts

Building models of financial markets is critical to all aspects of an ESG. Financial data informs the design of the models, provides benchmarks for calibration and serves as the ultimate frame of reference for model validation. The historical record of economic and financial markets is an indispensable guide to the dynamics that govern ESG model simulations. Detailed knowledge of these dynamics is essential for setting ESG model calibration targets and understanding strengths and weaknesses of various ESG model frameworks.

The chapter will familiarize the reader with the kinds of fundamental qualitative and quantitative features of economic and financial market variables that are common within ESGs. These important qualitative and quantitative features are commonly referred to as stylized facts.

Section 6.1 will begin with a definition of stylized facts and a discussion of the role that stylized facts play in understanding the strengths and weaknesses of various model frameworks and in setting calibration targets. Stylized facts refer to generalized interpretations of empirical findings that provide a basis for consistent understanding of markets or economic drivers across a wide range of instruments, markets and time periods. Analysis of historical data is commonly used as the basis for determining stylized facts and setting calibration targets; however, stylized facts can also be based on expert judgement.

Section 6.2 will run through some of the important considerations in determining the set of stylized facts that drive the calibration and validation of ESG models. Some of these considerations include the data window or span of time of the data being analyzed, and the usefulness of rolling averages and standard deviations. The stability of return statistics and correlations among asset classes also are important considerations.

Section 6.3 will provide an overview of historical characteristics of key economic and financial market variables that are commonly modeled within an ESG and will provide examples of stylized facts that practitioners may consider in their selection, calibration and validation of an ESG. In this, the role of expert judgement also will be briefly discussed. The section begins with a set of examples of typical stylized facts of financial market instruments and then addresses important considerations of interest rates, yield curves and various bond market characteristics as a way of further illustrating the development of stylized facts.

Section 6.4 provides additional examples using the equity markets, including the historical characteristics of return and volatility in these markets as contrasted with the bond markets. The section continues with a discussion of the leverage effect of market volatility on equity returns, leading to what is called fat-tailed behavior in the distribution of returns.

Section 6.6, an **appendix** to this chapter, reviews some sources of economic and financial market data.

6.1 STYLIZED FACTS

As practitioners seek a more in-depth understanding of market data, it is sometimes unsatisfying to discover that many apparently “intuitive” or “obvious” features about financial markets turn out to be anything but. Natural questions such as “What is the relationship between inflation and interest rates?” or “What is the relationship between the yield curve and future economic growth?” require detailed and nuanced answers. In fact, these sorts of questions are part of a large and active area of financial research and one for which the conclusions often shift.

While such questions are relevant for ESG practitioners, the answers to these questions are beyond the scope of this introductory chapter. The purpose of this chapter is to provide a context for setting stylized facts. It is up to the users of ESG models to determine the exact set of stylized facts that are appropriate for their specific application. Several references are provided in the annotated bibliography to guide the reader in pursuing more advanced considerations.

Stylized facts refer to generalized interpretations of empirical findings that provide a basis for consistent understanding of markets or economic drivers across a wide range of instruments, markets and time periods. Due to their generality, they are often qualitative. For example, one might employ as a stylized fact that the correlation between corporate bond returns and equity returns becomes larger (more positive) as the credit quality of the bonds

declines. This is a qualitative assumption that would have implications for the way in which the dynamics of corporate bonds and equities were modeled in the ESG. The specific setting of numerical values or targets for these correlations is also a stylized fact that would be of direct use in the calibration of the ESG.

Stylized facts are important in guiding the design of an ESG in that they help establish and prioritize the properties that the ESG model must have to be useful for a given application. They provide a basis for understanding the strengths and weaknesses of ESG model frameworks and the applicability of the underlying model equations to specific risk management applications. For example, ESG models for strategic asset allocation tend to be focused on means and volatilities, while economic capital models are more focused on tail events. The determination of stylized facts is an important part of the process of selecting an appropriate ESG model for a specific application.

Analysis of historical data is commonly used as the basis for determining stylized facts and setting calibration targets. Several sources can be used for economic and financial data. These include the Federal Reserve, Bloomberg, Ibbotson/Morningstar and Global Financial Data. However, stylized facts can also be based on expert judgment, as can the associated calibration targets. Expert judgment is the use of subjective and informed judgment by experienced professionals in the construction, calibration and parameterization of an ESG. For example, the current global interest rate environment has no clear historical parallel. Thus, one might adopt as a stylized fact that interest rates will remain low for an extended period of time. To fulfill this vision, an explicit calibration target would have to be based on expert judgment rather than on information that can be substantiated by the analysis of historical data.

6.2 CONSIDERATIONS IN THE DETERMINATION OF STYLIZED FACTS

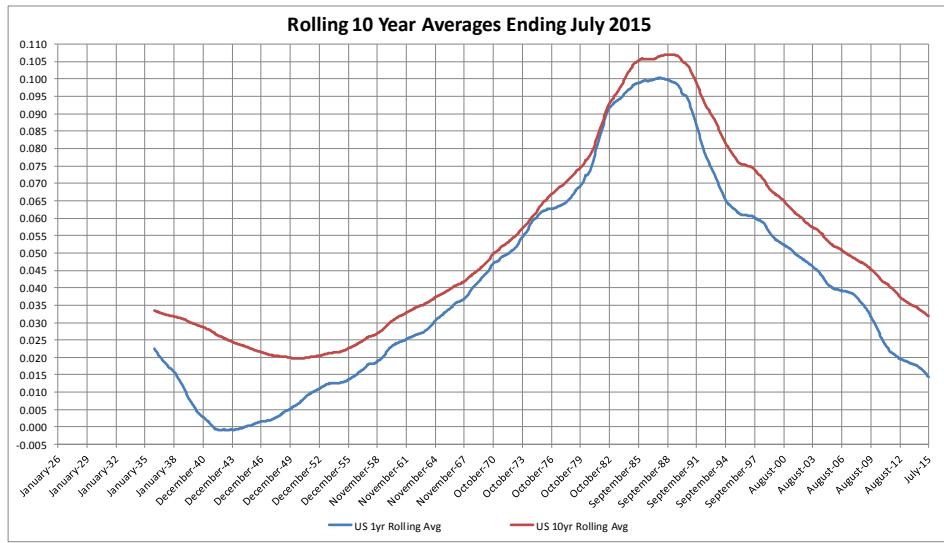
It is natural to summarize financial market variables in terms of their averages, standard deviations and correlations. These summary statistics tell a good bit of the story but do not inform the subtle but important aspects of how markets are experienced through time. More advanced applications such as those used for pricing and risk management typically require additional specifications that may include information related to distributional shapes (fat tails), pathwise behavior (how variables move over time) and the ways that characteristics of modeled variables change in different economic environments.

The model complexity and the number of parameters that make up the model equations will tend to increase as the need to incorporate more advanced stylized facts increases. Thus, the practitioner will need to evaluate the trade-offs between model complexity and the required dynamics of the modeled variables, to ensure that the ESG model is fit for the specific application.

6.2.1 DATA WINDOW CONSIDERATIONS

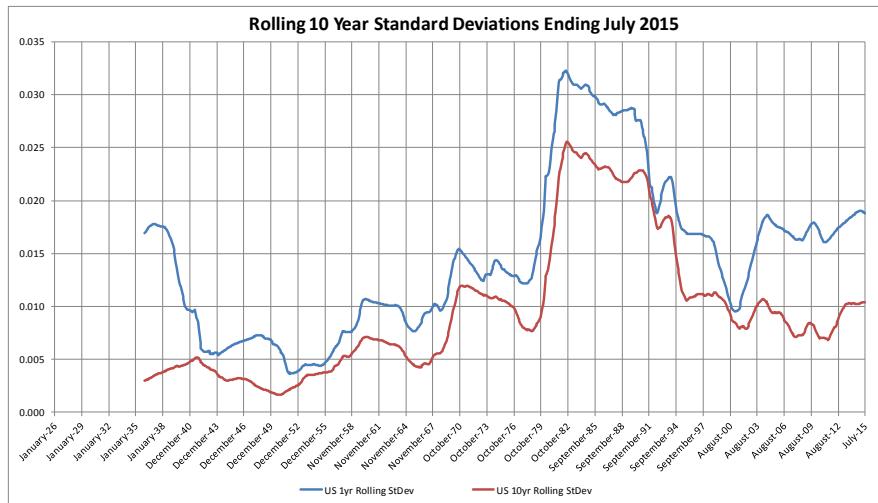
As an example of the need for more advanced stylized facts, it is illustrative to note how sensitive historical statistics can be to the data window for which they are calculated, and to get a sense of the possible range of parameter values for the variables being modeled. Rolling averages and standard deviations are often used for this purpose. Rolling statistics are computed at each observation time by looking back for a specific number of periods—for example, five years—and rolling this window forward as one advances through the observation times. Figures 6.1 and 6.2 show the rolling 10-year averages and standard deviations for the U.S. one-year and U.S. 10-year Treasury yields.

Figure 6.1
ROLLING 10-YEAR AVERAGES FOR U.S. 1-YEAR AND 10-YEAR TREASURY YIELDS



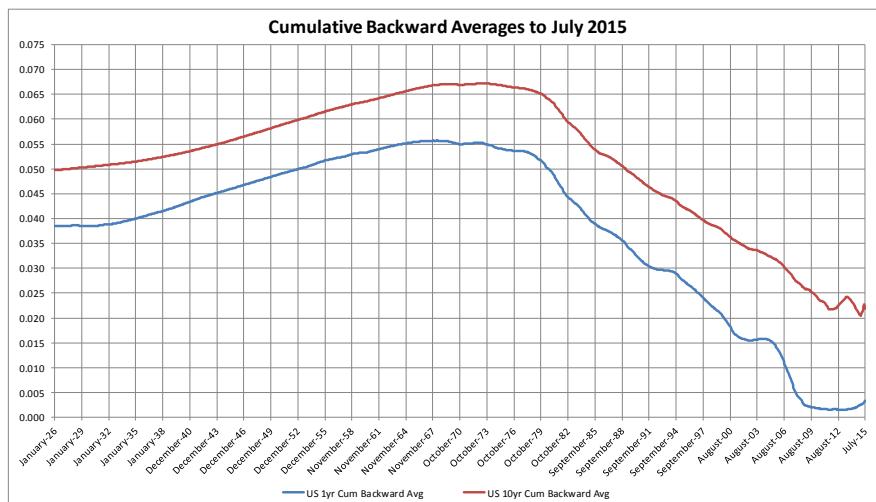
Sources: Ibbotson; Bloomberg; Conning, Inc.

Figure 6.2
ROLLING 10-YEAR STANDARD DEVIATIONS FOR U.S. 1-YEAR AND 10-YEAR TREASURY YIELDS

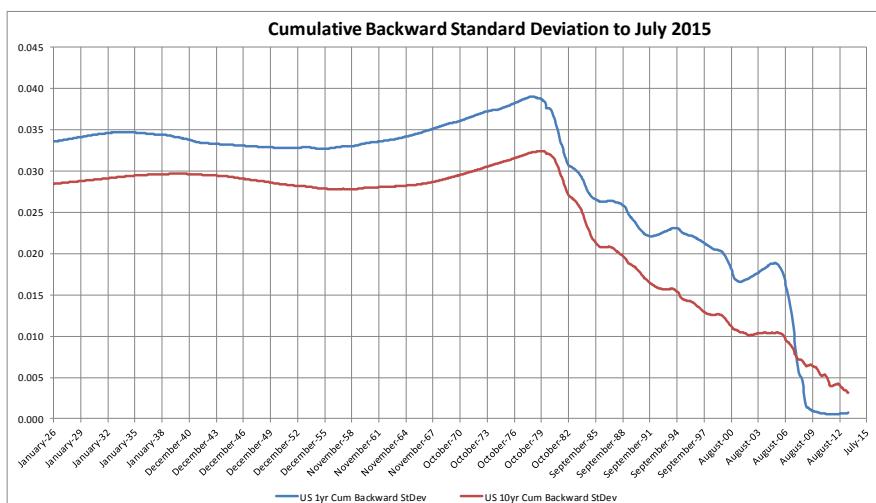


Sources: Ibbotson; Bloomberg; Conning, Inc.

Cumulative backward averages and cumulative backward standard deviations are another technique that is frequently used to observe changing mean and volatility levels over time. This technique computes statistics that show the change from the current observation time to the beginning observation time. As the beginning observation time advances, the variable statistics are computed over smaller and smaller data windows. Figures 6.3 and 6.4 show cumulative backward averages and cumulative backward standard deviations for one- and 10-year U.S. Treasury yields.

Figure 6.3**CUMULATIVE BACKWARD AVERAGES FOR U.S. 1-YEAR AND 10-YEAR TREASURY YIELDS**

Sources: Ibbotson; Bloomberg; Conning, Inc.

Figure 6.4**CUMULATIVE BACKWARD STANDARD DEVIATIONS FOR U.S. 1-YEAR AND 10-YEAR TREASURY YIELDS**

Sources: Ibbotson; Bloomberg; Conning, Inc.

As can be seen from these charts, even simple statistics such as means and standard deviations can vary substantially over time and over different historical time periods. Further, Figure 6.4 shows that over the most recent period, the 10-year yield has been substantially more volatile than the one-year yield. This is a rare occurrence and would likely place added tension on the ESG interest rate models that were designed and calibrated to reflect a more traditional relationship between long- and short-term bond yields.

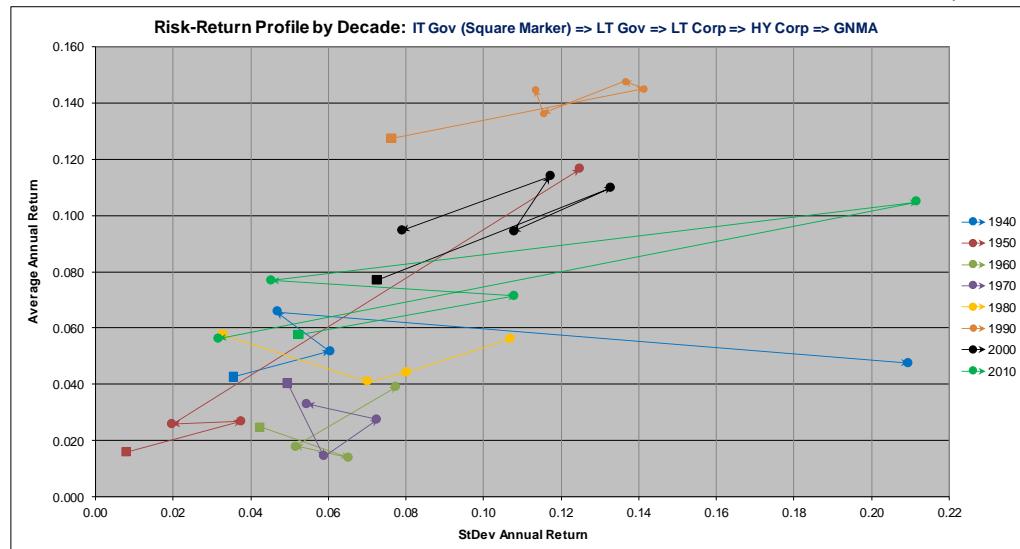
6.2.2 STABILITY OF RETURN STATISTICS AND CORRELATIONS

The relationship between different asset class returns changes over time, and these changes can occur for reasons that are difficult to predict. Figure 6.5 shows the relationships between risk and return for major fixed-income asset classes over the past eight decades. Each dot in the graph refers to an asset class, beginning with the square dot

(intermediate term government) and then connecting to long term government, then long-term corporate, then high yield corporate, and then GNMA. Each color refers to a different decade.

Figure 6.5

RELATIONSHIP BETWEEN RISK AND RETURN FOR MAJOR FIXED-INCOME ASSET CLASSES, 1940–2010

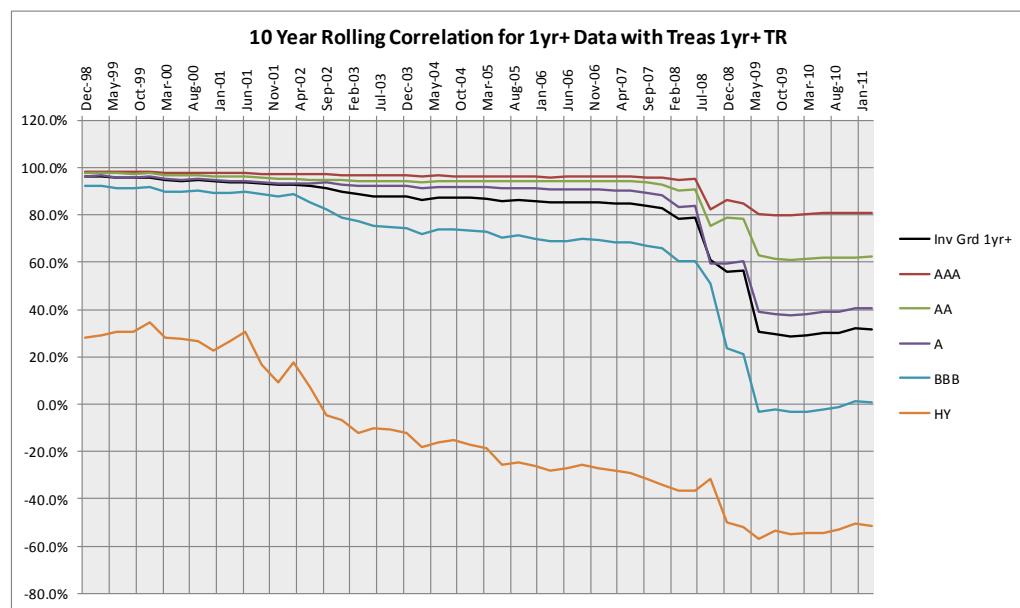


Sources: Ibbotson; Bloomberg; Conning, Inc.

The correlation between the returns on various asset classes also is highly variable. Figure 6.6 shows how correlations between corporate bonds and Treasury bonds have fluctuated over time.

Figure 6.6

CORRELATIONS BETWEEN CORPORATE BONDS AND TREASURY BONDS OVER TIME



Sources: Bloomberg; Conning, Inc.

Correlations can also vary depending on the frequency at which they are computed. For example, monthly return correlations can be quite different from quarterly return correlations, and there may be no clear explanation for the difference.

These changes to the characteristics and relationships between and among variables over time may be important to capture for some financial applications. For example, the pathwise nature of interest rates and equity market returns are important considerations in hedging risks associated with underwriting investment guarantees. An ESG that is being used for the management of variable-annuity risk would need to include model dynamics that capture the pathwise features of these variables. This consideration would guide the selection of the interest rate and equity components of an ESG that would be used for such an application.

6.3 OVERVIEW OF COMMONLY MODELED ECONOMIC AND FINANCIAL VARIABLES

Stylized facts are generalized interpretations of empirical findings that provide a basis for consistent understanding of markets or economic drivers across a wide range of instruments, markets and time periods. Due to their generality, they are often qualitative. Some examples of stylized facts might include the following:

- Yields for longer-maturity bonds tend to be greater than yields for shorter-maturity bonds.
- Monthly fluctuations in short-maturity yields tend to be larger than monthly fluctuations in longer-maturity yields.
- When short-maturity yields are low, longer-maturity yields are normally higher than the shorter-maturity yields (a normal shaped yield curve).
- When short-maturity yields are high, longer-maturity yields are often lower than shorter-maturity yields (an inverted yield curve).
- Interest rates can be negative.
- Corporate credit spreads are wider for lower-credit-quality instruments.
- Corporate credit spreads tend to fluctuate more during recessionary periods.
- The probabilities of default will fluctuate with general economic conditions and firm- or industry-specific conditions.
- Equity returns exhibit both higher expected returns and higher volatility than fixed-income returns.
- The volatility of equity returns fluctuates significantly.

Stylized facts typically come from an analysis of historical data and reflect the relevant considerations for the calculations that underlie a particular application. Thus, it is important for the practitioner to have a working knowledge of the economic and financial market variables being simulated by an ESG model. The following sections review general characteristics of some key economic and financial variables and demonstrate the types of analysis that might be used to develop a set of appropriate stylized facts.

6.3.1 INTEREST RATES AND YIELD CURVES

Interest rates are an important component of most ESG models, because they affect both the price of investments and the cost of liabilities on a company's balance sheet. In a simple sense, interest rates reflect the annualized cost of a specific financial obligation, generally expressed as a percentage of the principal of a loan and usually expressed on an annual basis. For an investor, the cost may translate to a yield or rate of return, based on the amount of principal (initial investment or price of the instrument) and the cash flow associated with interest payments and eventual return of principal of the bond as it matures. Interest rates may refer in a general sense to a "risk-free" loan such as reflected by a U.S. Treasury bond, or may refer to a specific loan where other factors (such as probability of default or lack of liquidity) may affect the interest rate.

Interest rates may vary in terms of time span of the obligation, over days, months or years. The relative differences in interest rates over a span of time, at any given point in time, for a specific bond market, are described in a construct known as a yield curve.

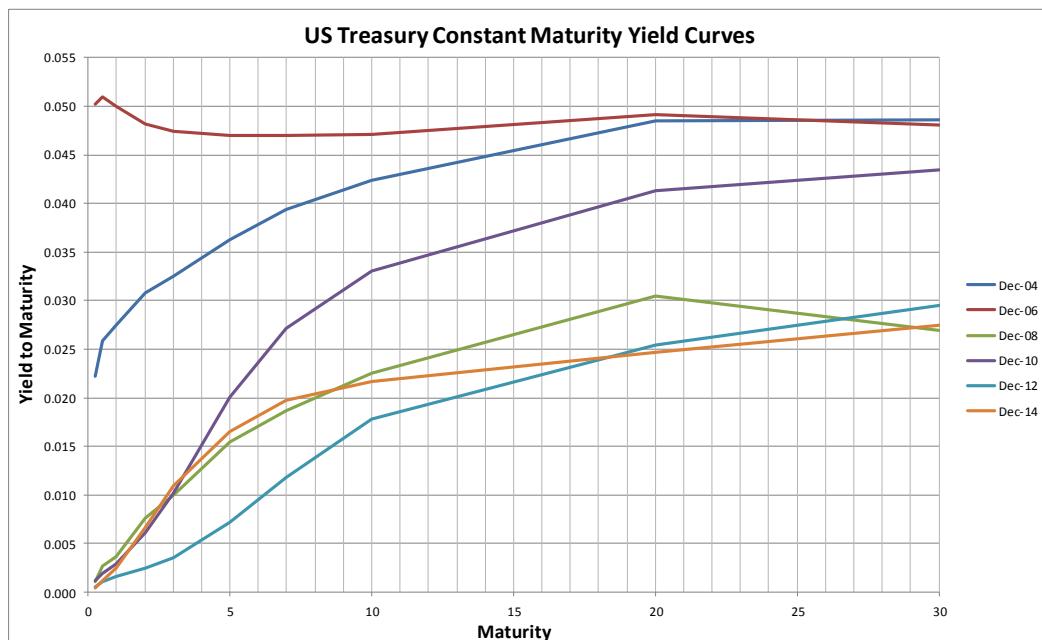
In a given bond market, there may be bonds with similar maturities but different yields to maturity. Information associated with individual bond securities must be compressed into a yield curve in a systematic fashion. In the end, yield curves are constructed objects based on traded prices of individual bonds or estimates of those prices. The

consolidation of yield information into yield curves is a convenient and concise way to summarize the status of a bond market.

6.3.2 Treasury Bonds

Bonds hold a relatively high position in the hierarchy of financial instruments in an ESG, because yields and returns on bonds are closely connected to interest rates. Risk-free bonds such as U.S. Treasuries embody a relatively pure representation of interest rates, because they are not complicated by significant concerns of liquidity or default. In the United States, Treasury bonds are issued with maturities up to 30 years. All bonds with original issue maturity in excess of one year pay semiannual coupons. All bonds with original issue maturity one year or less pay no coupon. The behavior of yields for shorter-maturity Treasury bonds is different from the behavior of yields for longer-maturity bonds. Figure 6.7 shows some actual U.S. Treasury yield curves from 2004 through 2014.

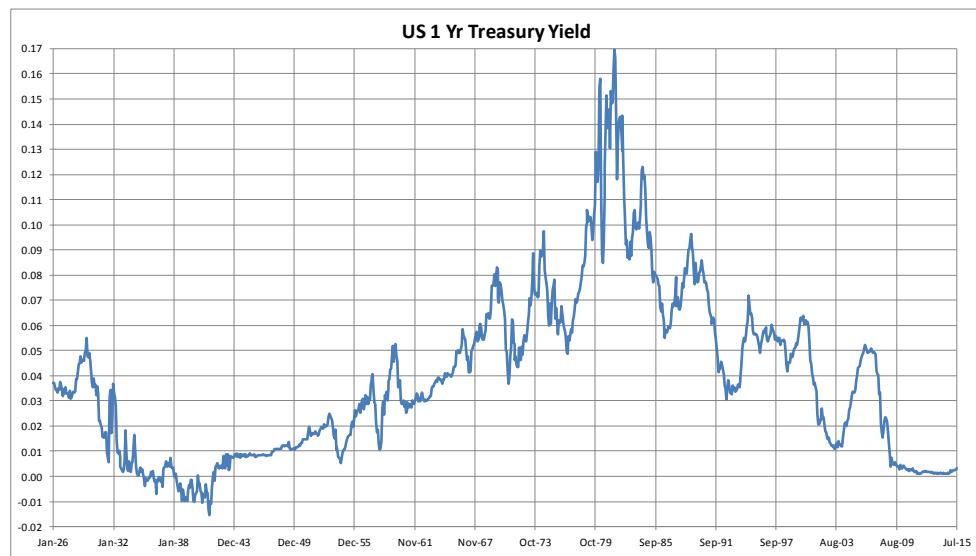
Figure 6.7
U.S. TREASURY YIELD CURVES, 2004–2014



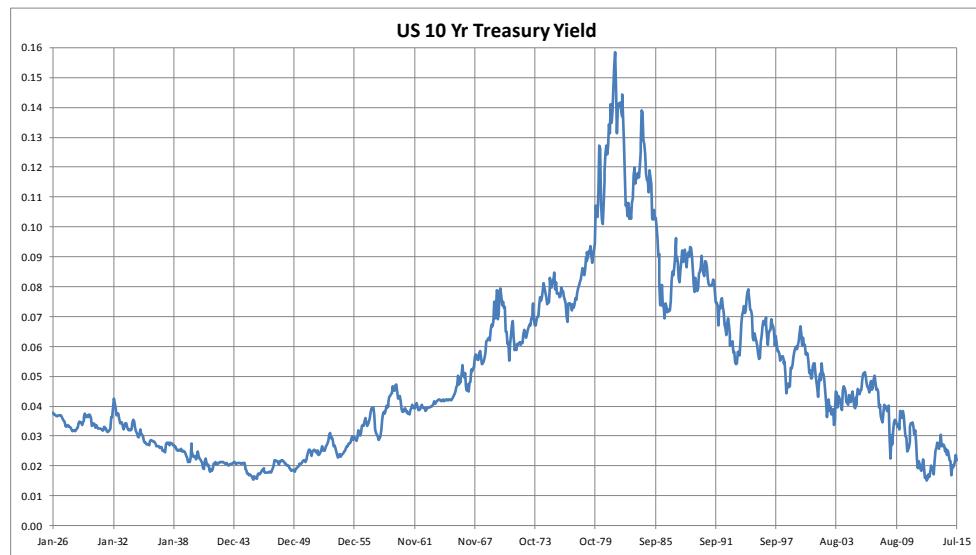
Sources: Bloomberg; Conning, Inc.

A casual review of this chart suggests that yields for longer maturities tend to be greater than yields for shorter maturities. The chart also shows that the very short end of the yield curve has been close to zero since 2008.

Figures 6.8 and 6.9 provide information about the historical levels of U.S. Treasury yields. Notice that the one-year yield was negative for a time during the 1930s and 1940s and has been very near zero but not negative since 2010. The 10-year yield in recent years has hit lows not seen since 1945.

Figure 6.8**HISTORICAL LEVELS OF 1-YEAR U.S. TREASURY YIELDS**

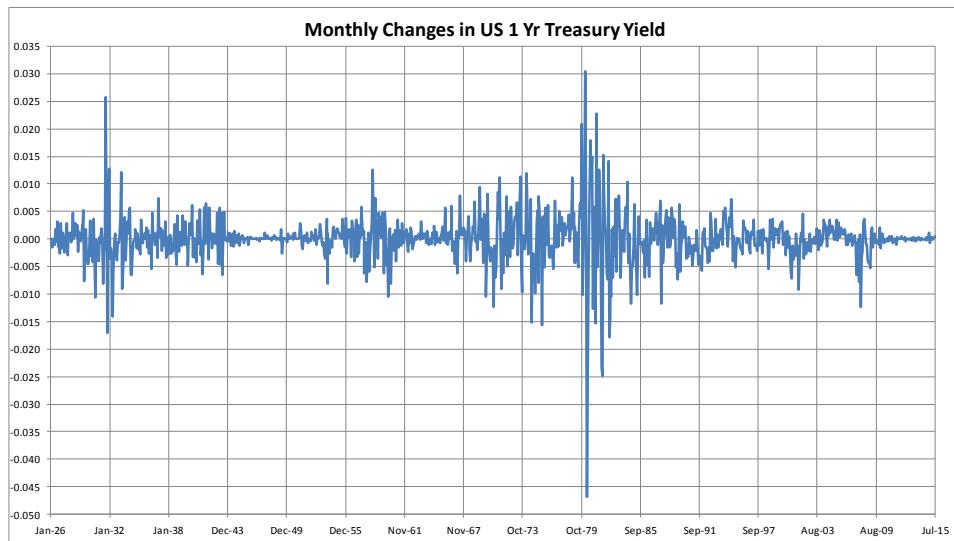
Sources: Ibbotson; Bloomberg; Conning, Inc.

Figure 6.9**HISTORICAL LEVELS OF 10-YEAR U.S. TREASURY YIELDS**

Sources: Ibbotson; Bloomberg; Conning, Inc.

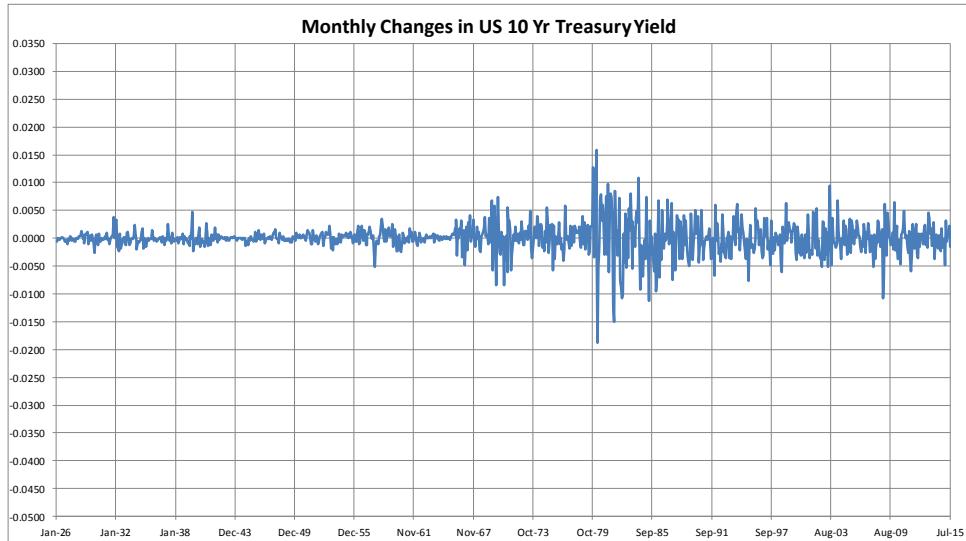
Monthly changes in yields are shown in Figures 6.10 and 6.11. We see that monthly fluctuations in the one-year yield tend to be larger than monthly fluctuations in the 10-year yield. We also observe that in recent years, the 10-year yield has fluctuated more than the one-year yield. Lastly, we note that the monthly fluctuations in the 10-year yield are larger from the late 1960s to present than they were prior to the late 1960s.

Figure 6.10
HISTORICAL MONTHLY CHANGES IN 1-YEAR U.S. TREASURY YIELDS



Sources: Ibbotson; Bloomberg; Conning, Inc.

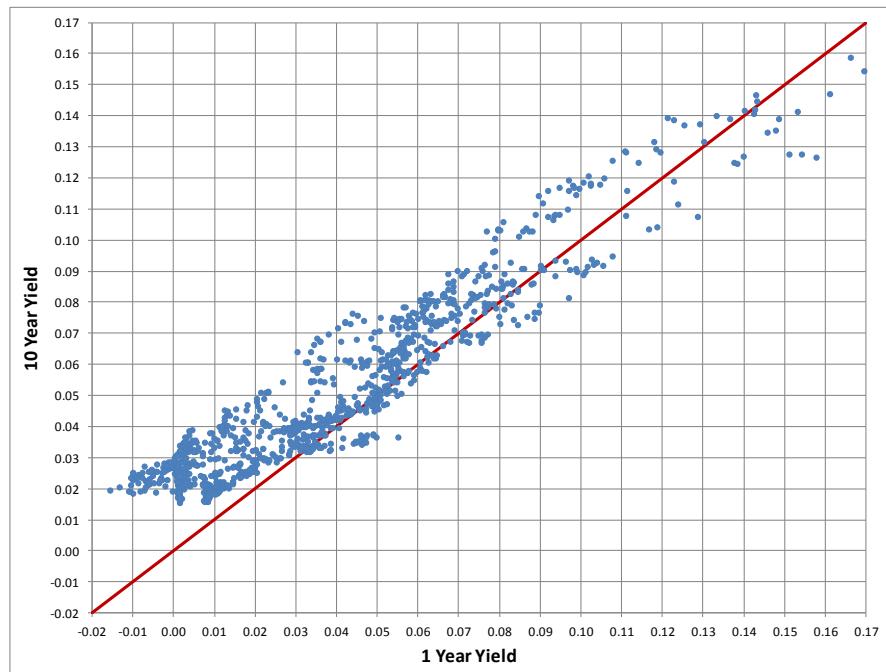
Figure 6.11
HISTORICAL MONTHLY CHANGES IN 10-YEAR U.S. TREASURY YIELDS



Sources: Ibbotson; Bloomberg; Conning, Inc.

Figure 6.12 shows that short- and long-term Treasury yields are highly correlated but there is some possibility for the relative positions of the one-year and 10-year yields to vary. The red line shown on the chart is a 45-degree line passing through the origin, which is useful for visually determining whether the one-year yield is less than or larger than its 10-year yield counterpart. When the one-year yield is low, the 10-year yield is normally larger than the one-year yield—a normal shaped yield curve. When the one-year yield is high, the 10-year yield is often smaller than the one-year yield—an inverted yield curve. On occasion, inverted yield curves have occurred when the one-year yield is at a relatively low level, as we saw for the December 2006 yield curve in Figure 6.7.

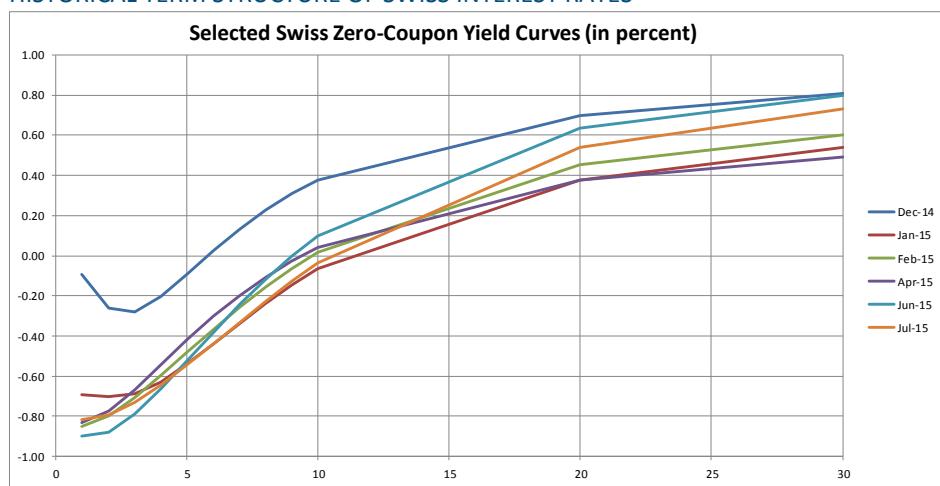
Figure 6.12
HISTORICAL CORRELATION BETWEEN 1-YEAR AND 10-YEAR U.S. TREASURY YIELDS



Sources: Ibbotson; Bloomberg; Conning, Inc.

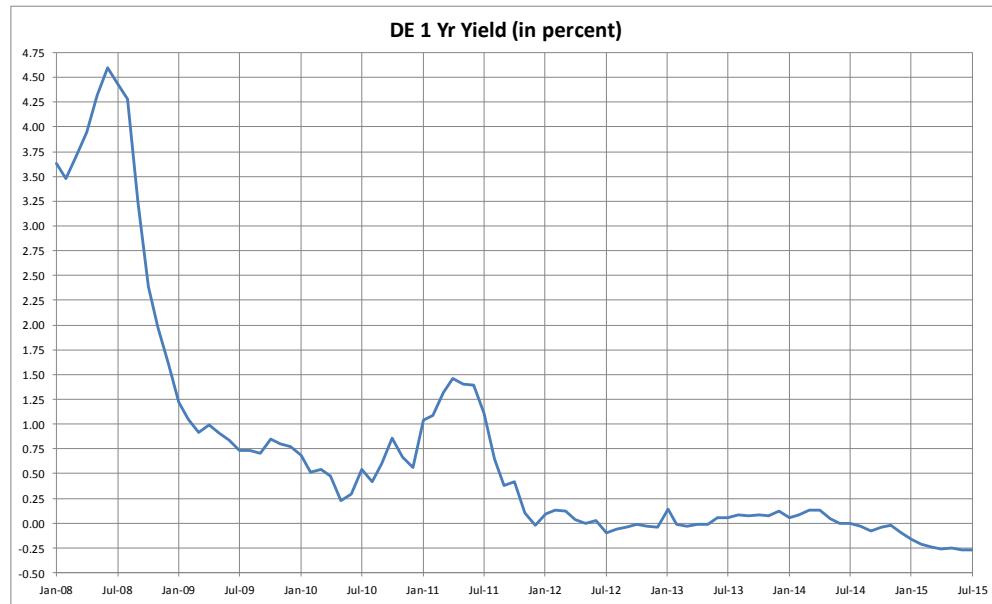
The question of negative interest rates has recently become relevant. Some central banks in Europe have set negative funds rates, thereby pushing short-term yields negative. At the same time, aggressive quantitative easing programs have driven up the price of longer-term bonds to such an extent that their yields are close to zero or even negative. Figure 6.13 shows the recent term structure of interest rates as computed and provided by the Swiss National Bank. It is interesting to note that long-term discount factors out as far as 10 years have been negative. And in Figure 6.14, the recent trajectory for the German one-year interest rate shows a rapid collapse in yields, with the one-year yield ultimately settling at a negative value.

Figure 6.13
HISTORICAL TERM STRUCTURE OF SWISS INTEREST RATES



Source: Swiss National Bank,

Figure 6.14
HISTORICAL 1-YEAR YIELD IN GERMANY



Source: Bundesbank,
http://www.bundesbank.de/Navigation/EN/Statistics/Time_series_databases/Macro_economic_time_series/macro_economic_time_series_node.html

There is also evidence of very small negative interest rates at the short end of the yield curve occurring in the United States after the financial crisis. Recent economic data suggests that negative interest rates do occur, and this is something the ESG user needs to consider in assessing and validating ESG interest rate projections.

6.3.3 CORPORATE BOND MARKET

For capturing the risk dynamics inherent in the corporate credit markets, three elements are important to consider. These include (1) default, along with the resultant recovery rates; (2) changing credit spreads; and (3) credit migration. This section will discuss each of these components in the context of historical observation.

6.3.4 DEFAULT AND RECOVERY

A corporate bond is a security with a fixed schedule of payments, generally consisting of a scheduled maturity date and associated coupons and a redemption amount. The receipt of the scheduled payments by the bondholder is contingent upon the issuing corporation being able and willing to pay them as scheduled. In contrast to Treasury bonds, corporate bonds involve a significant possibility that payments may not be made as scheduled. If the payment schedule is broken, default is said to have occurred.

The probabilities of default will fluctuate with general economic conditions and firm- or industry-specific conditions. A wide range of outcomes are possible once a bond enters default, ranging from a total loss of all remaining scheduled payments to relatively small losses or an extension of the payment schedule. The amount that the owners of a defaulted bond receive is called the recovery amount. A given bond either will survive to its maturity date without migrating to default or will migrate to default prior to its scheduled maturity date.

6.3.5 RATINGS

Most corporate bonds will be rated during their lives. These ratings are provided by an independent rating agency, and the purpose of the rating is to capture the likelihood that the bond will make its scheduled payments. The rating for an individual bond will tend to fluctuate over time. The manner in which the rating fluctuates is called rating

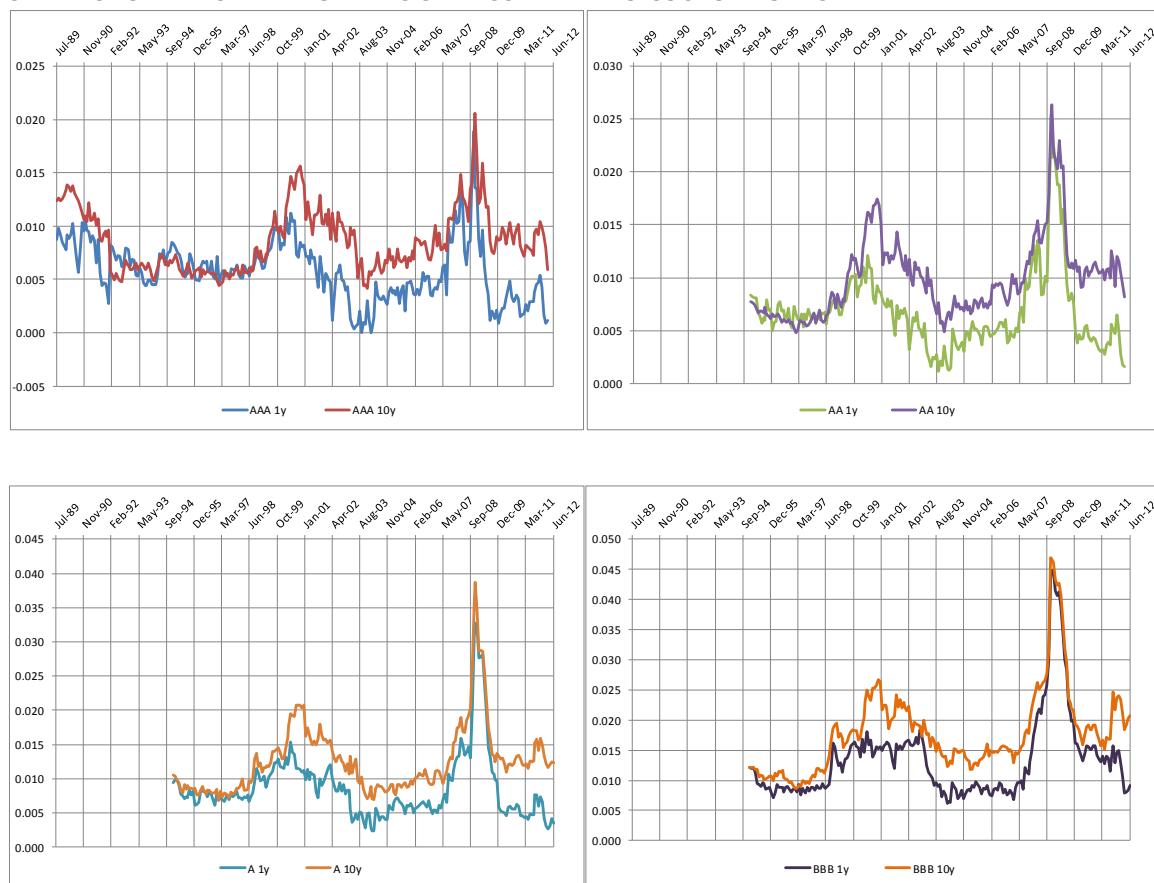
migration. Investment-grade bonds are those bonds with an S&P rating of BBB or higher. Non-investment-grade or "high-yield" bonds are those with an S&P rating below BBB.

6.3.6 YIELDS AND CREDIT SPREADS

Using a method similar to that for Treasury bonds, a yield to maturity for corporate bonds can be readily computed. Subject to the issues that were discussed for Treasury bonds, one can also produce a yield curve for the various rating classes in the corporate bond market. Since corporate bonds have risk that a comparable Treasury bond does not have, the yield on a corporate bond has been historically higher than the yield on a comparable Treasury. For a given maturity, the amount by which the corporate bond yield exceeds the Treasury bond yield is called the spread.

Spreads are a convenient way to summarize the current state of the corporate bond market. Figure 6.15 shows spreads for investment-grade U.S. industrial zero-coupon bonds at one-year and 10-year maturities.

Figure 6.15
SPREADS FOR INVESTMENT-GRADE U.S. INDUSTRIAL ZERO-COUPON BONDS



Sources: Bloomberg; Conning, Inc.

One of the most striking features of the data is the extraordinary widening of spreads at the peak of the 2008 financial crisis. Generally, we see that spreads are wider for lower credit quality. Spreads are also wider for longer-term bonds of the same credit quality.

Figure 6.16 shows monthly changes in U.S. spreads. Recession dating from the National Bureau of Economic Research (NBER) has been superimposed.

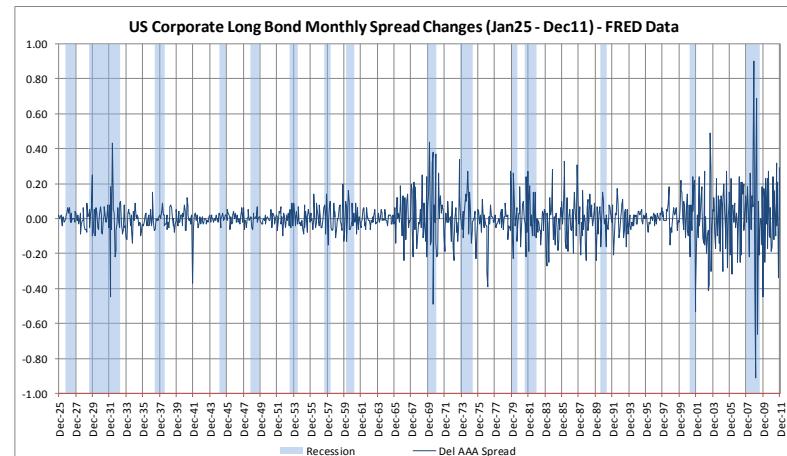
Figure 6.16
MONTHLY CHANGES IN U.S. SPREADS, 1989–2012



Sources: Bloomberg; Conning, Inc.

There is a tendency for spreads to fluctuate more during recessionary periods, but the relationship is not definitive. The relationship is more pronounced with a longer data series, as Figure 6.17 shows. This spread series clearly exhibits different behavior in and out of recession. Historical data indicates that credit costs represent only a fraction of the spread on corporate bonds, and this suggests that some portion of corporate bond spreads is due to factors other than credit costs (such as liquidity.) More on this will be presented in Chapter 12.

Figure 6.17
MONTHLY CHANGES IN U.S. SPREADS, JANUARY 1925–DECEMBER 2011



Sources: Moody's; Bloomberg; Conning, Inc.

6.3.7 CREDIT TRANSITION AND MIGRATION

A credit transition matrix specifies the probability of a bond remaining at its current credit rating (the elements on the main diagonal) and the probabilities of moving to another credit rating, including default. Figure 6.18 shows an illustrative credit transition matrix. The illustrative matrix is for a one-year transition, which is the typical frequency used in benchmarking an ESG. The numbers in this transition matrix clearly indicate that, on average, there is a stronger tendency for a bond in each rating class to downgrade than to upgrade. This is typical of real-market data.

Figure 6.18

ILLUSTRATIVE CREDIT TRANSITION MATRIX

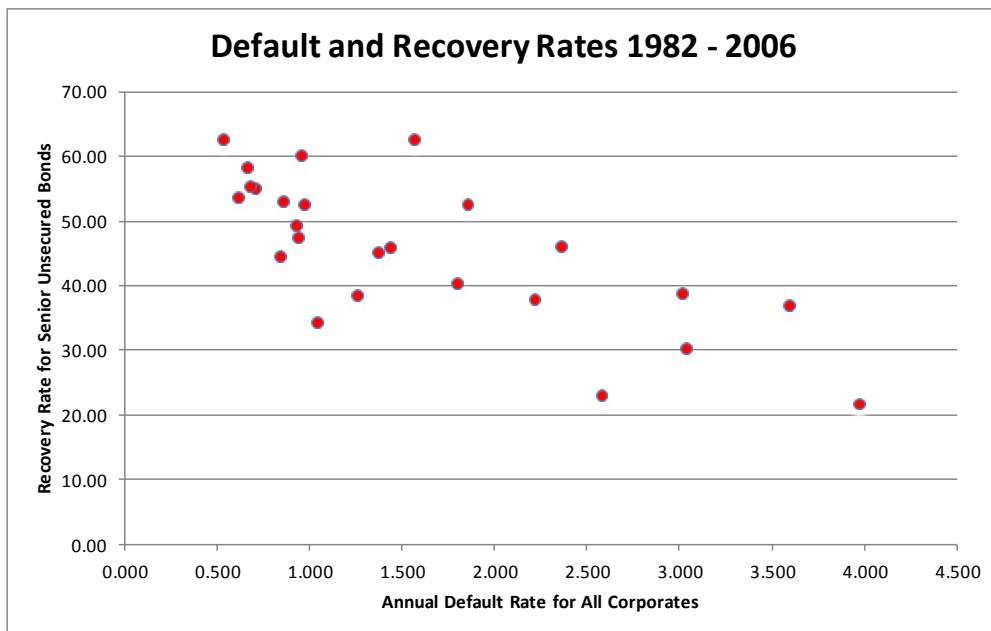
Final	Aaa	Aa	A	Baa	<Baa	Default
Aaa	0.8951	0.0850	0.0116	0.0049	0.0036	0.0000
Aa	0.0152	0.8848	0.0796	0.0125	0.0073	0.0007
A	0.0008	0.0318	0.8869	0.0586	0.0148	0.0072
Baa	0.0004	0.0028	0.0477	0.8683	0.0683	0.0126
<Baa	0.0001	0.0005	0.0015	0.0120	0.9307	0.0553

Sources: Moody's; Bloomberg; Conning, Inc.

To the extent that a corporate bond enters into default, the bondholder is normally able to get a recovery settlement. This settlement, referred to as the recovery rate, is typically quoted as a percentage of the initial par value of the bond. Figure 6.19 shows the probability of default and the average recovery rate on those defaulted bonds over the period 1982–2006.

Figure 6.19

DEFAULT AND RECOVERY RATES, 1982–2006



Source: Moody's; Conning Analysis

6.4 OVERVIEW OF THE EQUITY MARKETS

The characteristics of equity investments differ significantly from those of fixed-income investments. The most obvious differences are that equities have historically exhibited both higher expected returns and higher volatility than fixed-income returns. These can be considered important stylized facts of the equity markets.

On an annual return basis, for the period 1926 through 2014, U.S. large-cap equities experienced an average return of 12% with a 20% standard deviation. By comparison, for the same period, also on an annual basis, U.S. long-term high-grade corporate bonds experienced an average return of 6.4% with an 8.4% standard deviation.

At a higher observation frequency—for example, monthly data—a more detailed examination of equity return data shows that the volatility of equity returns fluctuates significantly. The technical term for this phenomenon is heteroscedasticity. In an ESG context, one can capture this effect by using models that include stochastic volatility processes, and ESG modelers will often refer to equity returns as exhibiting stochastic volatility.

A frequently considered stylized fact about equity returns is that they exhibit a leverage effect, which means that when markets become more volatile, the risk of a large sudden decline in equity prices becomes elevated. This can be described in the context of fat-tailed behavior of the distribution of returns. There are significant implications for the ESG concerning the manner in which equity returns are experienced on a path-by-path basis if such a stylized fact is adopted. Indeed, an ESG that is calibrated to such a stylized fact would produce some simulated paths with pronounced market crashes, which would increase the left tail of the return distribution and make it more expensive to hedge equity return risk.

6.5 CHAPTER SUMMARY

1. Stylized facts refer to generalized interpretations of empirical findings that provide a basis for consistent understanding of markets or economic drivers (across a wide range of instruments, markets and time periods). Due to their generality, they are often qualitative.
2. Stylized facts help with assessing the strengths and weaknesses of ESG model frameworks and the applicability of the underlying model equations to specific risk management applications.
3. Stylized facts are also used in setting of calibration targets that drive the model calibration and validation processes.
4. Historical data is often used as the basis for determining stylized facts and setting calibration targets.
5. Several sources can be used for economic and financial data. These include the Federal Reserve, Bloomberg, Ibbotson/Morningstar and Global Financial Data.
6. Stylized facts and related calibration targets typically reflect facts such as these:
 - a. Means and standard deviations can vary considerably over different historical periods.
 - b. The dynamics of interest rates at very low, average and high levels differ in terms of simple measures such as the volatility of changes.
 - c. Negative interest rates can occur.
 - d. Real interest rates fluctuate significantly over time.
 - e. The risk and return attributes of the major asset classes can change over time.
 - f. Corporate bond spreads fluctuate significantly over time.
 - g. Widening corporate spreads do not always track realized default rates.
 - h. Credit costs represent only a fraction of the spread on corporate bonds, and this suggests that some portion of corporate bond spreads is due to factors other than credit costs (e.g., liquidity).
 - i. There is a correspondence between the movement of corporate bond spreads, bond migrations and defaults with recessions, but this relationship is not precise.
 - j. Equity return distributions exhibit fat-tailed behavior.
 - k. Volatility fluctuations and default clustering occur across time; one does not experience average behavior.
 - l. Correlations between modeled economic and financial market variables are not stable over time and can depend on whether one is using monthly, quarterly or annual observations.
7. Recent economic conditions are causing a reassessment of where a natural level for interest rates might be and what history can tell us about what lies in the near future. Historical benchmarks can be distorted by

recent events, and a significant amount of expert judgment may be needed in arriving at a set of stylized facts for financial markets.

6.6 APPENDIX: DATA SOURCES

ESG practitioners are well served by having a wide range of data sources at their disposal. Historical data is the standard for setting stylized facts and targets that can be used for ESG model calibration and validation.

Economic and financial market data may be free or require a subscription. All other things being equal, it is desirable to use data available at a higher frequency. ESG practitioners typically require data at monthly or quarterly frequency. Some applications call for daily frequency, but that is not typical. The frequency at which data is available can vary. Financial data such as interest rates or return series are generally available at monthly frequency for recent time periods. GDP data is released quarterly. Less current financial data might be available only at annual frequency.

If an ESG is part of an ongoing risk management program, the usual requirements are that the data source be accurate, regularly updated and capable of automated download. This is because in this use context, the ESG is likely to be recalibrated, validated and initialized to market conditions at regular intervals such as year-end or quarter-end.

The following commercial data vendors offer an extensive collection of financial data amenable to electronic download across a broad range of financial variables and economies:

- Bloomberg
- Thomson-Reuters
- Global Financial Data

Morningstar provides a range of popular data, including the Ibbotson SBBI Classic Yearbook and the Dimson, Marsh, Staunton Global Returns Data.

Many central banks and other government organizations provide economic data. Listed here are examples of these data sources:

- St. Louis Fed (Federal Reserve Economic Data)
- Board of Governors of the Federal Reserve System
- Bank of Canada
- Statistics Canada (CANSIM)
- Bank of England
- Ministry of Finance (Japan)
- Swiss National Bank
- German Bundesbank
- U.S. Bureau of Labor Statistics
- U.S. Bureau of Economic Analysis

There is a considerable amount of useful data scattered in the academic literature. Such data may appear in journal articles, in books or on websites.

Bond-rating services, such as Moody's, Standard & Poor's, Fitch and DBRS, provide periodic reports containing data useful for corporate bond models. This data includes rating transitions, default probabilities, recovery rates and combined credit costs.

6.6.1 AD HOC DATA SOURCES

Canadian Institute of Actuaries. *Report on Canadian Economic Statistics, Annual Data File*.

Dimson, Elroy, Paul Marsh, and Mike Staunton. *Triumph of the Optimists: 101 Years of Global Investment Returns*. Princeton, NJ: Princeton University Press, 2002.

Eitrheim, Øyvind, Jan T. Klovland, and Jan F. Qvigstad, eds. 2004. *Historical Monetary Statistics for Norway, 1819–2003*. Occasional Paper no. 35. Oslo: Norges Bank.

—, eds. 2007. *Historical Monetary Statistics for Norway, Part II*. Occasional Paper no. 38. Oslo: Norges Bank.

Shiller, Robert. n.d. “Online Data.” Yale University, <http://www.econ.yale.edu/~shiller/data.htm>.

6.7 CHAPTER REFERENCES

Campbell, J., A. Lo, and C. MacKinlay. 1997. *The Econometrics of Financial Markets*. Princeton, NJ: Princeton University Press.

Cont, R. 2001. Empirical Properties of Asset Returns: Stylized Facts and Statistical Issues. *Quantitative Finance* 1: 223–236.

Singleton, K. 2006. *Empirical Dynamic Asset Pricing: Model Specification and Econometric Assessment*. Princeton, NJ: Princeton University Press.

Chapter 7: Model Calibration and Parameterization

ESG Users need to incorporate a view of future market dynamics into their risk-modeling environment. Will interest rates fall or become negative in the next five years? What is the probability of a global credit crisis in the next 12 months? What will be the value of the large-cap equity index tomorrow? If these questions were to be independently posed to a number of investment and risk professionals, the responses would likely result in a range of answers. Each would form a view based on history, current information regarding the state of the economy, and expert judgment.

The process of reflecting these views into an ESG is referred to as model calibration. Different model calibrations may be required for different applications such as the following:

- **Regulatory risk management.** Some elements of regulatory regimes, such as Article 226 (Use Test) of the European Insurance and Occupational Pensions Authority (EIOPA) Solvency II Directive, require firms to take into account the assumptions of their internal model in making business decisions. Having ownership of the underlying assumptions of the economic models is a key element in satisfying such regulatory requirements.
- **Stress testing.** Stress testing is a widely used tool in risk management. Stressing the core assumptions of models represents an alternative view that allows for the impact of extreme and unexpected events to be studied and quantified.
- **Strategic asset allocation (SAA).** SAA is commonly based on analysis of expected returns and risk from an ESG. The SAA calibration may embed the shorter-term assumptions of the internal model into a multiyear stochastic projection, and/or it may reflect a long-term steady-state view of the economy and financial markets.

To effectively perform any of these (and other) tasks, one must have consistent and justifiable views across different parts of the business and have the people and processes in place to implement these views within the company's ESG model. It is this implementation step that often proves to be the most challenging for users of statistical models such as those found within an ESG.

Section 7.1 of this chapter outlines the calibration process and introduces the notion of a cascade structure. Calibration is the process of setting the parameters of the equations within an ESG model. It is usually performed as part of an iterative process that is repeated with parameter adjustments until a reasonable fit relative to the needs of the application is achieved.

Section 7.2 discusses calibration as it relates to market-consistent ESG models. Model calibration in this context is based on how well the model reproduces the market values of the universe of the traded derivatives used to calibrate the model. For nontraded securities or instruments, pricing would then be determined from the prices of similar traded securities (e.g., minimum crediting rates proxied by a series of very long dated puts).

Section 7.3 discusses calibration as it relates to real-world ESG models. Parameterizations of real-world ESG models require the user to make choices about the future economic environment that he or she wants to reflect in the risk analysis work, including appropriate steady-state levels, appropriate values for the initial conditions, the key parameterization targets, the expected reversionary paths of economic variables, and so forth.

Section 7.4 discusses two methods (maximum-likelihood estimation and Kalman filtering) that can be used to help with the calibration of more sophisticated real-world ESG model equations.

7.1 THE CALIBRATION PROCESS

Calibration is the process of setting the parameters of the equations within an ESG model to produce the distributions and dynamics (e.g., volatility, correlations, tail characteristics) of economic and financial variables required by the application for which they are being used. The calibration process associated with different models will depend on the sophistication of the equations that underlie the models. A mean-variance model has parameters that are

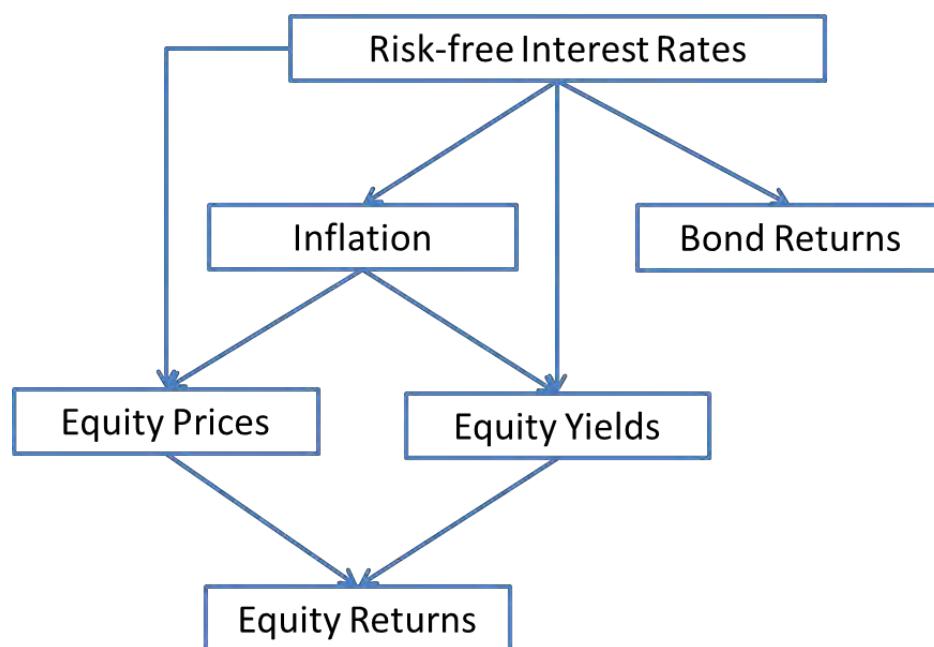
identical to the assumptions that drive these models—namely mean, standard deviation and correlations. This makes the calibration of these models pretty straightforward.

As the model equations increase in complexity, the number of model parameters increases, and the linkage between the input parameters and the desired output statistics become less intuitive. The calibration process thus becomes increasingly challenging as additional variables are incorporated into the model. To make the calibration process practical, many ESG models make use of what is referred to as a cascade structure.

A cascade structure is a model framework design whereby each subsequent variable in the cascade structure depends only on prior values of the variable and the values of variables that lie above them on the cascade structure. The variables at the top of the cascade are calibrated against an established set of calibration criteria. The variables in the lower levels of the cascade structure are subsequently calibrated against their own calibration criteria, factoring in the appropriate relationship to higher-level variables. There is no feedback loop mechanism within the cascade structure, which means that once the variables that reside in the top levels of the cascade are appropriately calibrated, the calibration of lower-level variables will not affect these previously calibrated variables.

The sequence of the cascade is not intended to reflect any cause-and-effect relationship but rather is designed to ease the calibration process while allowing variables to be appropriately correlated. For example, making use of a cascade structure where interest rates are at the top of the cascade with inflation below would not imply that interest rate changes cause inflation. It would, however, allow the simulation of interest rates and inflation to have an appropriate relationship, so that when interest rates are high, inflation will tend to be high, and vice versa. Figure 7.1 shows a cascade structure that might be used within an economic scenario generator.

Figure 7.1
CASCADE STRUCTURE OF A HYPOTHETICAL ECONOMIC SCENARIO GENERATOR



The calibration processes for real-world and market-consistent ESG models are similar but differ significantly in the criteria used to determine goodness of fit. Thus, they warrant separate discussions.

7.2 CALIBRATIONS OF MARKET-CONSISTENT ESG MODELS

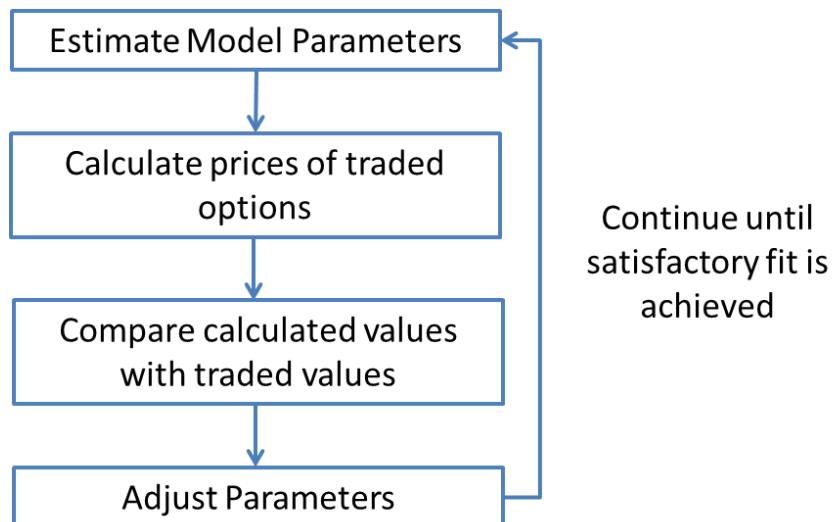
Market-consistent valuation applications require ESGs to be capable of generating scenarios that can reproduce the observable prices of traded derivative instruments. ESGs that are used for these purposes need to adhere to strict

mathematical properties that are designed to satisfy risk-neutral and arbitrage-free conditions. These same scenarios are then used to determine prices for derivative instruments and insurance contracts with embedded options that are not traded but require market valuation. The process of parameterizing an ESG to reproduce observable prices of traded derivative instruments is referred to as “model calibration.”

Because the model calibration process is designed to reproduce the prices of traded derivatives, the ultimate calibration depends on both the pricing date and the set of traded derivatives used to calibrate the model. The validation associated with the model calibration is based on how well the model reproduces the market values of the universe of traded derivatives used to calibrate the model. The calibration process associated with a market-consistent ESG model is shown in Figure 7.2.

Figure 7.2

MARKET-CONSISTENT CALIBRATION PROCESS



The idea behind market-consistent model calibration is to generate scenarios whereby the cash flows from a financial product or instrument can be appropriately discounted so as to provide an estimate of the current market price of the underlying product or instrument. Because market-consistent models are designed to be risk-neutral, all financial instruments will have the same expected return, and that expected return will be equal to the risk-free rate. The risk-free rate can thus be used to discount the cash flows from a product or financial instrument to generate its price.

To ensure that the prices estimated for nontraded securities are reasonable, the model is first calibrated to reproduce prices of traded securities, including options that are in line with actual observable market prices. The first step in this calibration process is to select a set of appropriate traded financial instruments for which current prices are readily available. Ideally, they would be similar in form to the nontraded products or instruments being priced (e.g., minimum crediting rates look like a series of very long dated puts).

Once the traded instruments have been identified, a source for the pricing information needs to be found. There are two common sources, private and public, for obtaining this information. Private data usually comes from broker's tear sheets. There are several key advantages to using this source data. First, since the prices are determined as a group, they tend to be internally consistent. This can be particularly important when attempting to calibrate the ESG model to these prices. Second, this is the way that companies typically value the options on their balance sheet. Thus, the use of these prices can create a tighter link between asset and liability valuations. The one drawback of private data is access: tear sheets are usually available only to active option traders. When this type of data isn't readily available, companies typically turn to actual trade data from a source such as Bloomberg. In this case, care must be taken to avoid using thinly traded securities, since their prices tend to be less reliable.

The calibration process shown in Figure 7.2 starts with a set of estimated model parameters. Using these parameters, the model is run, and the scenarios generated are used to calculate prices for a set of traded options. The prices

generated under the model scenarios are compared with the actual traded values to determine goodness of fit. If the calculated option prices replicate the actual traded prices of the options within an acceptable tolerance, the parameters are deemed to be acceptable. If not, the model parameters are re-estimated, and the process is repeated until an appropriate fit is found.

7.3 PARAMETERIZATION (CALIBRATION) OF REAL-WORLD ESG MODELS

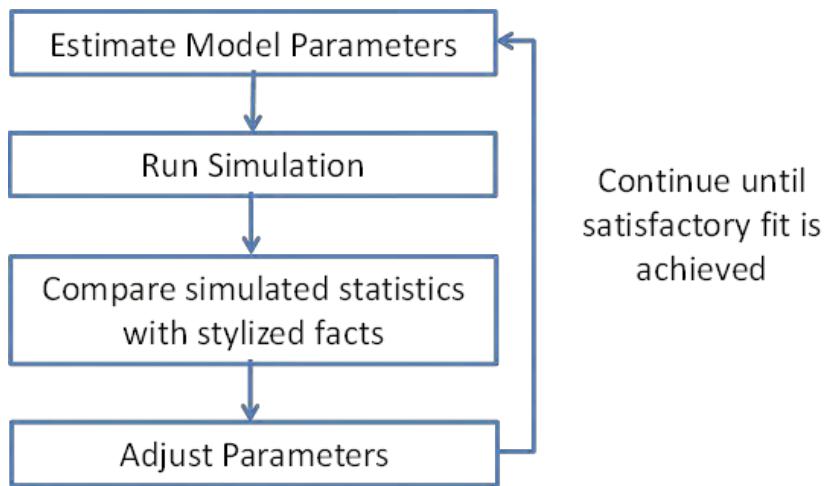
To distinguish real-world calibrations from market-consistent calibrations, we will refer to the real-world calibration process as “model parameterization.” Most risk management applications, for example, require ESGs to be capable of producing dynamics (e.g., volatility, correlations) that are representative of the possible future paths of economic variables. Commonly referred to as “real-world” calibrations, they answer managers’ what-if questions as the managers try to gauge the likelihood of future events and the impact on their business.

Because real-world parameterizations are forward looking, they require explicit views as to how the economy will develop in the future, so they require a significant amount of expert judgment to determine the veracity of the scenarios that result from the parameterization process. In practice, real-world calibrations often are parameterized to be consistent with historical dynamics of economic variables, although the long-term steady-state levels associated with these parameterizations can differ from long-term historical averages in favor of current consensus expectations or even individual viewpoints.

Parameterizations of real-world ESG models require users to make choices about the future economic environment they want to reflect in their risk analysis work. Some of the key decision points when parameterizing a real-world model include (a) selecting the appropriate steady-state levels, (b) determining the appropriate values for the initial conditions, (c) identifying the key parameterization targets or “stylized facts” that are necessary for the application, (d) controlling the expected reversionary paths of economic variables, and (e) considering general assumptions.

The parameterization process for a single economic or financial variable follows the flow chart depicted in Figure 7.3.

Figure 7.3
REAL-WORLD PARAMETERIZATION PROCESS

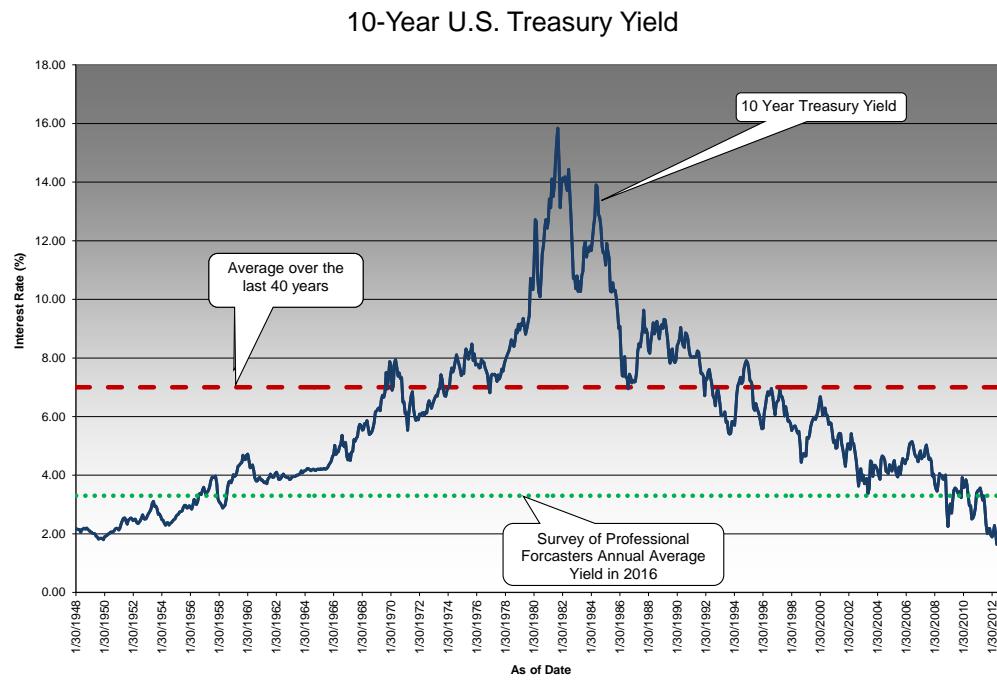


7.3.1 SELECTING THE APPROPRIATE STEADY-STATE LEVELS

Steady-state levels are values that define the long-term expected mean levels of economic or capital market variables. One can set these steady-state levels to reflect historical averages over some specified time period, consensus long-term economic forecasts, or a particular company or economic viewpoint. Figure 7.4 shows the type of information that might be used to help set the assumed steady-state level of 10-year U.S. Treasury yields within a company’s ESG model parameterization. The figure shows that 10-year U.S. Treasury yields have ranged between 2% and 16% since 1948 and that the average yield over that 40-year period has been 7.0%. In addition, the Philadelphia

Federal Reserve's Survey of Professional Forecasters indicates that the annual average for the 10-year yield in 2016 was expected to be 3.3%. This variability demonstrates the challenges the model user faces in selecting appropriate assumptions for risk management applications. The steady-state level that is ultimately selected will need to involve expert judgment and will depend on the specific application for which the model is being built.

Figure 7.4
10-YEAR U.S. TREASURY YIELD



Source: Federal Reserve Board, Philadelphia Federal Reserve's Survey of Professional Forecasters

7.3.2 DETERMINING THE APPROPRIATE VALUES FOR THE INITIAL CONDITIONS

Depending on the particular application, it may be appropriate to set the initial conditions consistent with a particular historical date (e.g., for financial valuation), a set of hypothetical starting conditions (e.g., for sensitivity testing), or at their steady-state levels (e.g., for SAA). The third option is useful to explore risk management strategies that do not depend on an assumption of an upward or downward bias in key economic metrics, such as capital losses on bonds resulting from rising interest rate projections.

7.3.3 IDENTIFYING THE KEY PARAMETERIZATION TARGETS OR "STYLIZED FACTS" NECESSARY FOR THE APPLICATION

ESGs need to balance complexity and practicality. The historical data can generate an infinite number of statistics. To keep the models academically credible yet appropriate for practical application, it is necessary to limit the factors used to describe the underlying dynamics of the modeled economic variables. As a result, it is not feasible for these models to hit every parameterization target or capture every stylized fact (historically observed relationship). Expert judgment thus is used to determine which targets are most important for an intended application. Short-term tail events may be deemed more critical for one-year regulatory capital calculations, whereas long-term mean and standard deviation measures may be of greater importance for strategic investment analysis. The importance placed on these different parameterization targets can have a material impact on the final model calibration and its applicability for specific risk management applications.

7.3.4 CONTROLLING THE EXPECTED REVERSIONARY PATHS OF ECONOMIC VARIABLES

Given a set of initial conditions and expected long-term steady-state levels, a real-world ESG model will generate a set of scenarios that move from initial levels to their steady-state levels over time. These mean reversion dynamics are often calibrated to be consistent with historical reversionary dynamics. While this default reversionary movement is appropriate for many risk management applications, some users of ESG models may want to reflect expected movements that differ from those implied by the modeled mean reversion parameterization. As an example, users may want the expected movement of economic variables to follow some specific internal or external economic forecast. To the extent that such a specific economic forecast is desired, the model will need to be capable of overriding the default pathwise behavior with some user-specified path. This is typically accomplished by shifting the default scenarios in such a way as to appropriately conform to a user-specified path. Alternatively, one could attempt to change the internal model parameters to change the default pathwise behavior, but given the complexity of the underlying ESG model equations, this could be a difficult task for many casual ESG users.

7.3.5 GENERAL ASSUMPTION CONSIDERATIONS

The considerations of choice for steady-state levels, initial conditions, stylized facts and reversionary paths demonstrate that, when one is setting economic and capital market assumptions, the appropriate scenarios to use for a particular analysis will depend on what a company is trying to accomplish.

- Scenarios used for regulatory capital assessments and stress testing require particular attention to extreme events; focus thus tends to be on the tails of the distributions.
- Business planning requires consistency with initial economic and capital market conditions, along with a forecast of the expected future path of the economy that is consistent with internal company viewpoints.
- Strategic risk analysis, which includes strategic asset allocation, more often concentrates on risk as it relates to the entire distribution of outcomes and replaces specific individual company viewpoints with a steady-state view of the economy and financial markets that reflects a neutral or consensus perspective.

7.4 METHODS OF PARAMETERIZING REAL-WORLD ESGS

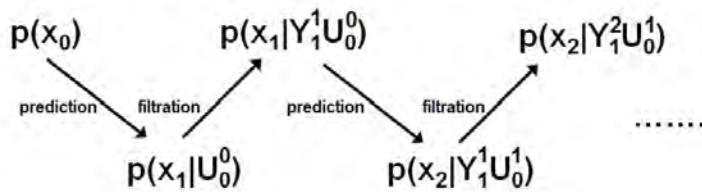
As the parameterization (calibration) process from Figure 7.3 implies, parameterization is an iterative process. Given the sheer number of parameters involved, a trial-and-error process would generally be impractical. Even many of the standard actuarial techniques, such as simple regression, are not effective, because the historical data for many economic variables violates one of the prime assumptions of these techniques: independent observations—as, for example, with interest rates and inflation, where tomorrow's level is highly dependent on today's observation. As a result, most parameterizations turn to more robust tools in this step.

The first such tool is maximum-likelihood estimation (MLE). The MLE process starts by calculating a likelihood function based on a possible set of model parameters $\{p_i\}$ and the observed data $\{x_j\}$. The result ends up looking like $\text{Prob}(x_j | \{p_i\} \text{ and } x_1, \dots, x_n)$. The inner calculation reflects the probability that the model would produce the x_j th observation given the selected parameter set and all the prior observations. Once that function is determined, the set of parameters that maximize MLE can be determined. The likelihood function is then the product of these values across all the historical observations. For computational ease, this process often uses the equivalent approach of maximizing the log of this function.

The second key tool that can assist with parameterization is Kalman filtering. Kalman filtering is designed to deal with several problems that crop up in real-world parameterizations. First, historical data has both errors and uncertainties in the process. These result from such issues as measurement noise and systems errors (e.g., in financial markets from liquidity effects). Second, with the increasing complexity of ESG models, there is an increasing reliance on non-observable state variables. While this makes these models more powerful, it adds an extra layer of uncertainty to the parameterization process that needs to be addressed. In practice, Kalman filtering dynamics result from the consecutive cycles of predicting the state of an observed variable based on a model, comparing that prediction with the realized outcome in the historical observed data, and updating the parameters to achieve optimal predictive power.

The Kalman filter works by propagating the conditional probability distribution function (PDF) $p(x_k | Y_1^k, U_0^{k-1})$, in other words by transitioning from $p(x_k | Y_1^k, U_0^{k-1})$ to $p(x_{k+1} | Y_1^{k+1}, U_0^k)$. In practice, this is done in two steps. First is a prediction step, in which the PDF at $k + 1$ is predicted with the information available at k —i.e., the conditional PDF $p(x_{k+1} | Y_1^k, U_0^k)$ is predicted. In the second step, the filtration, the Kalman filter updates the information on x_{k+1} to optimally improve the state estimate after making the observation Y_{k+1} . The Kalman filter performs successive prediction and filtration cycles for all k (see Figure 7.5).

Figure 7.5
THE PREDICTION AND FILTERING CYCLE OF THE KALMAN FILTER



7.5 CHAPTER SUMMARY

- Calibration is the process of setting the coefficients and parameters of the equations within an ESG model to produce the distributions and dynamics required by the application for which they are being used.
- The calibration process associated with different models will depend on the sophistication of the equations that underlie the models.
- A cascade structure is a model framework design whereby each subsequent variable in the cascade structure depends only on prior values of the variable and the values of variables that lie above them on the cascade structure.
- The ESG model calibration is a function of the application for which the resulting ESG scenarios will be used to inform.
- Similar processes are used for the calibration/parameterization of market-consistent and real-world ESG models, but the processes differ significantly in the criteria used to determine goodness of fit.
- Market-consistent valuation applications require ESGs to be capable of generating scenarios that can reproduce the observable prices of traded derivative instruments. Because the model calibration process is designed to reproduce the prices of traded derivatives, the ultimate calibration depends on both the pricing date and the set of traded derivatives used to calibrate the model.
- Parameterizations of real-world ESG models require that users make choices about the future economic environment they want to reflect in their risk analysis work. Some of the key decision points when parameterizing a real-world model are (a) selecting the appropriate steady-state levels, (b) determining the appropriate values for the initial conditions, (c) identifying the key parameterization targets or “stylized facts” that are necessary for the application, (d) controlling the expected reversionary paths of economic variables, and (e) considering general assumptions.
- While historical data plays a significant role in most real-world ESG model calibrations, many areas require the use of expert judgment in the calibration process.
- Maximum-likelihood estimation and Kalman filtering are examples of methods that can be used to assist in the real-world ESG model calibration process.

7.6 CHAPTER REFERENCES

- Bickel, P. J., and K. A. Doksum. 1997. *Mathematical Statistics*. Oakland, CA: Holden-day Inc.
- DeGroot, M. H., and M. J. Schervish. 2002. *Probability and Statistics*. 3rd ed. Boston: Addison-Wesley.
- Federal Reserve Bank of Philadelphia. n.d. Survey of Professional Forecasters.
<https://www.philadelphiafed.org/research-and-data/real-time-center/survey-of-professional-forecasters/>.
- Gelb, A. 1974. *Applied Optimal Estimation*. Cambridge, MA: MIT Press.

- Harvey, A. C. 1990. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.
- Hull, John. 2006. *Options, Futures, and Other Derivative Securities*. Englewood Cliffs, N.J.: Prentice Hall.
- Jazwinski, Andrew H. 1970. *Stochastic Processes and Filtering: Mathematics in Science and Engineering*. New York: Academic Press.
- Kalman, R. E. 1960. A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering* 82(1): 35-45.
- Kalman, R. E., and R. S. Bucy. 1961. *New Results in Linear Filtering and Prediction Theory*. . *Journal of Basic Engineering* 83(1): 95-108
- Liu, W., J. C. Principe, and S. Haykin. 2010. *Kernel Adaptive Filtering: A Comprehensive Introduction*. New York: Wiley.
- Maybeck, Peter S. 1979. *Stochastic Models, Estimation, and Control*. Mathematics in Science and Engineering 141-1. New York: Academic Press.
- Simon, D. 2006. *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*. New York: Wiley-Interscience.
- Spanos, A. 1999. *Probability Theory and Statistical Inference*. Cambridge: Cambridge University Press.

Chapter 8: Model Validation

Validation ensures that the estimation of an ESG's parameters result in simulated behavior that is a good representation of the variable or market under consideration. Effective validation of an ESG requires comparing simulated output data to some predefined benchmark and acceptance criteria. Benchmarks may be based upon historical data, internal views, regulator demands or a combination of these factors. Chapter 6 discussed the model specification process and the development of stylized facts that can be derived from historical empirical financial data. Chapter 7 went into detail regarding the process of calibrating the model. This chapter will introduce important considerations when validating the output of an economic scenario generator (ESG). Validation of real-world and market-consistent ESGs is discussed separately.

Section 8.1 will describe general considerations in developing a validation system, including the importance of automated validation approaches, establishment of acceptance criteria and data considerations. Acceptance criteria should be based on what the end user expects of the model, and would be defined before any data are analyzed or validation performed. This is preferable to validation based on problem discovery, which attempts to discover problems with the output after the model is run. Data considerations address issues of accuracy, completeness and appropriateness of the data, in terms of what the end user expects out of the model.

Section 8.2 addresses issues specific to validation of real-world ESGs, beginning with checking that initial conditions of the model match those of the market (point-in-time validation). Real world model validation can involve a great many quantitative and qualitative possibilities for determining the quality of a scenario set, so considerable expert judgement is required. This section also explores principles of in-sample and out-of-sample validation, and explores a further description of the types of variable output that should be subject to validation. In-sample refers to using data that is consistent with what was used to estimate model parameters. Out-of-sample validation means back-testing the model to a historical point in time and seeing how it performs looking forward. Validation should not be restricted to the primary variable to be simulated, but should also include derived variables, correlations and interdependencies in order to get a fuller picture of the quality of the ESG.

Section 8.3 will describe issues specific to validation of market-consistent ESGs. Validation of market-consistent ESGs is more straightforward than for real-world ESGs: it only needs to be validated for market consistency and risk neutrality. Market consistency involves determining the quality of fit of the output against market pricing, using a goodness of fit statistic or applying qualitative judgement. Validating risk neutrality confirms that the expected return on a particular asset class is the same as the expected return on a risk free asset, which is accomplished through application of a martingale test.

Section 8.4 gives some final thoughts regarding effective validation and introduces a flow chart that shows the main steps of an idealized validation process.

8.1 GENERAL CONSIDERATIONS IN DEVELOPING A VALIDATION SYSTEM

For a typical insurance or pension undertaking, the list of financial and economic variables that may be of interest is typically quite large. For this reason, the validation system and environment requires careful design at inception in order to organize the various data elements in an ordered fashion.

An automated validation system is preferable to manual validation. Validation should be repeatable and consistent through time. Automated validation may include the production of goodness of fit statistics, which can aid the reviewer in making decisions. However, it should be noted that sole reliance on such statistics may lead to false conclusions and missed problems, due to the limitations of the statistical measure itself and the relevance and completeness of the available market data. For this reason, effective validation will invariably need a human to visually review and make informed observations about the results. The professionals doing this should ideally have knowledge of the models, the market being considered, the end use of the model output and the features of the model at prior dates in order to most effectively validate the model.

8.1.1 ACCEPTANCE CRITERIA

Before any data are analyzed or validation performed, it is helpful to form the acceptance criteria upon which the model output will be judged. These acceptance criteria should be based on what the end user expects the model to do. An idealized process for forming acceptance criteria might be as follows:

1. Select a person or persons to formulate acceptance criteria. Ideally, this would be a group made up of the direct users of the system, the end user of the scenarios or derived data, participants in the market to be validated and risk model experts, as well as individuals who are independent of the system usage to provide oversight.
2. Decide which economic variables are to be validated, and determine the materiality of these variables.
3. Formulate concrete acceptance criteria, which should be based on a combination of analysis of market data, expert judgment and an understanding of the sensitivities to and materiality of particular risk factors of a firm. Acceptance criteria should not be arbitrary but instead justifiable and based on data analysis and informed judgment.
4. Define when a model is accepted and when rejected. This is usually best dealt with by scoring the ESG output against particular acceptance criteria and holistically considering the extent to which it matches all the acceptance criteria. For all but the most simplistic uses of an ESG, it is likely that some areas will perform better than other areas; therefore, it is better to answer the question "How well does the ESG as a whole perform?" than to reject a model because a single acceptance criterion is not adequately met.

This type of approach to validation, whereby the particular desirable features of an ESG are based on analysis of a firm's risk exposures, is preferable to what might be called a "problem discovery" approach. In a problem discovery approach, a user first runs the ESG, creating a large output data set, and then tries to discover problems with the output. Because all models are to some extent a misspecified simplification of reality, it is highly likely that this approach will yield a list of potential issues. The problem with this approach is twofold:

1. It becomes difficult to objectively and in an unbiased fashion determine if any of the issues are truly material or if they would have appeared on a list of acceptance criteria had they been formed prior to validating the output.
2. It focuses only on what is wrong with the model and not on the many other things that may be right with the model. Thus, it is insufficient in providing a holistic view of the model quality.

Therefore, it is recommended that the validation process start with acceptance criteria and then move on to the validation stage. Chapter 6 discussed the model specification process and the development of stylized facts that form the basis of the acceptance criteria.

With the acceptance criteria in place, the next stage is to actually validate the ESG and determine its appropriateness to the application for which it is intended. Usually, validation entails comparing the output of the ESG with market data, and finally with the acceptance criteria, which may be based on market data or a combination of market data and expert judgment. In this process, there are several considerations to take into account.

8.1.2 DATA CONSIDERATIONS

Adequately validating a model requires that appropriate data be collected, stored and maintained. This requires consideration of the data source to be used, the historical window against which to compare the ESG output, the frequency of data available to be used, and the quality of the data. In particular, for thinking about the validation data, issues of accuracy, completeness and appropriateness are central to the discussion.

Accuracy. Accurate refers to the degree of confidence that can be placed in the data. To be accurate means the data is reliable and appropriately reflects the behavior of the economy or asset class. Data must be sufficiently accurate to avoid material distortion of the model output. Accurate data does not contain any material errors or omissions. The data should also be recorded and stored both in a timely fashion and consistently over time.

The assessment of accuracy consists of monitoring in areas such as the initial generation of data, along with how the data is stored and processed, to ensure errors are not inadvertently introduced into the data.

Completeness. Complete means that databases provide comprehensive information. More specifically the data shall, where possible, meet the following conditions:

- Be available for all material asset classes and all relevant model variables
- Be of sufficient granularity so that the trends and the full understanding of the behavior of the underlying risks can be identified and understood
- Take account of all information that may be relevant to the understanding and modeling of risk

Appropriateness. Appropriate means the data does not contain biases that make it unfit for its purpose. The data should be relevant to the model. Relevance in this aspect means that the data wherever possible should satisfy these requirements:

- Directly relates to the underlying risk drivers
- Provides a good guide for the future when it is used to calculate projections
- Applies to the current risks and time point

Taking these considerations into account ensures that the data is fit for the purpose for which it will be used.

In reality, it is not always possible to source data that is accurate, complete and appropriate for every data item that one may wish to analyze. It is therefore important at the outset to define what to do in these circumstances. This typically involves extrapolation from the available historical data and/or from other economies.

8.2 SPECIFIC CONSIDERATIONS IN VALIDATION OF REAL-WORLD ESGS

Validation of real-world ESGs against market data falls into three distinct categories: point-in-time validation, ex-post or in-sample validation, and ex-ante or out-of-sample validation.

8.2.1 VALIDATION CATEGORIES

Point-in-time validation. Point-in-time validation involves checking that the time zero characteristics or initial conditions of the model adequately match those of the market. For instance, one should be sure that the initial yield curves of the market and the model are the same before a simulation is run.

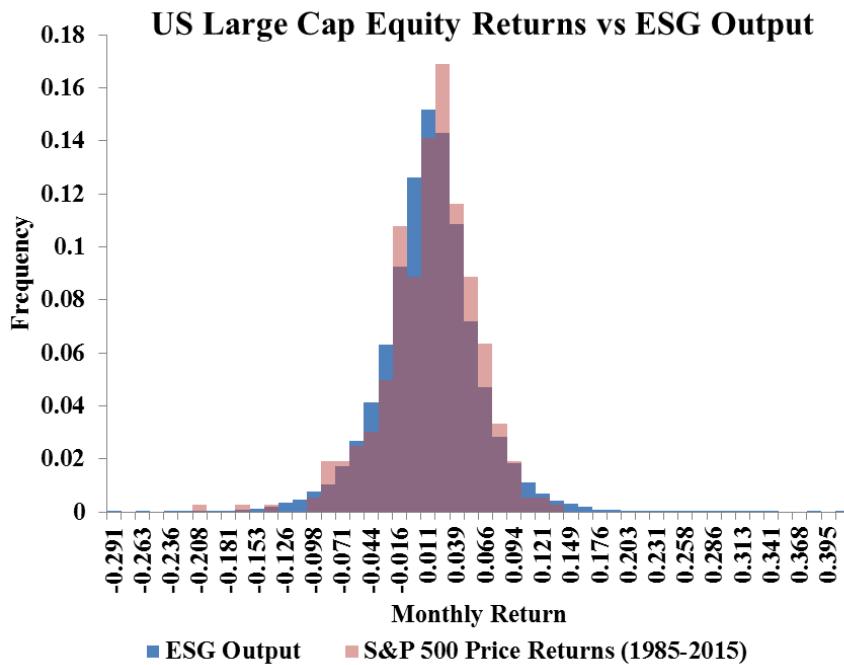
Ex-post or in-sample validation. In-sample validation involves taking historical data and determining the extent to which the stylized facts of the data are captured by the ESG. The comparison is typically based on the same data period that was used for the estimation of the model parameters; hence it is referred to as in-sample. This type of validation may take many forms but is likely to include the following tasks:

- A visual comparison of the distributions or probability density functions (PDFs) of the market data and ESG output, such as that shown in Figure 8.1
- A comparison of key statistics (e.g. mean, median, skewness and kurtosis)
- A comparison of correlations or bivariate distributions

We should, however, be careful when drawing conclusions from these in-sample comparisons. In model building, it is not usually the case that the best approach is to blindly fit the model to market data. Such models may pass statistical goodness of fit tests; however, they may perform poorly ex ante.

Figure 8.1 shows a comparison of the probability density function of S&P 500 price returns between 1985 and 2015 with that of a real-world ESG model calibrated to this data.

Figure 8.1
ACTUAL VERSUS SIMULATED DISTRIBUTIONS OF S&P 500 PRICE RETURNS



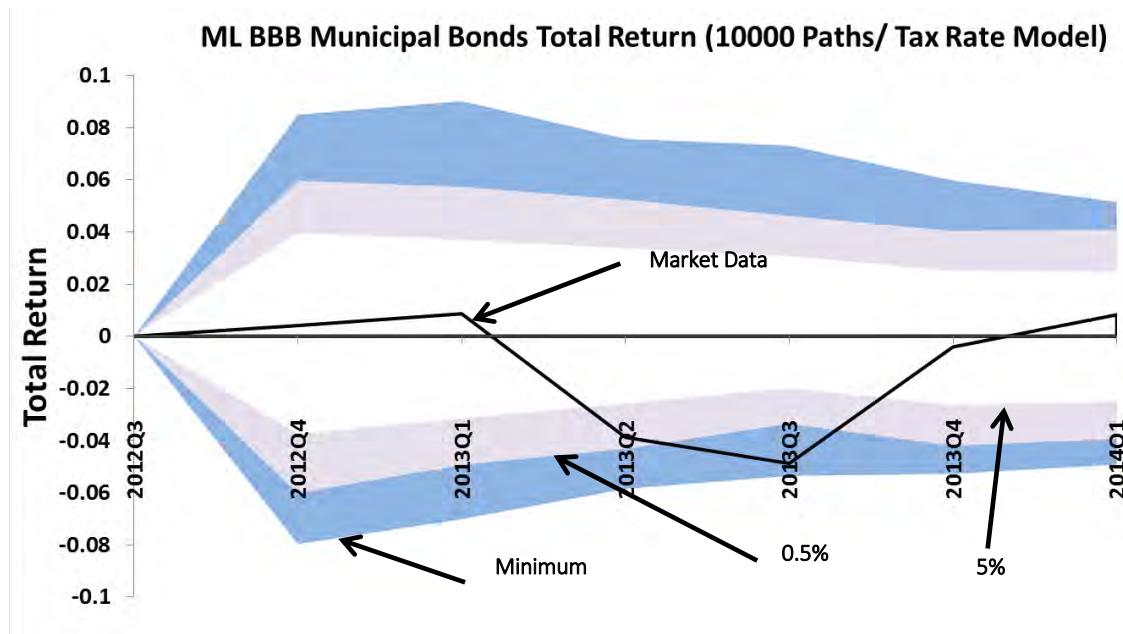
Sources: Bloomberg; Conning, Inc.

Ex-ante or out-of-sample validation. Out-of-sample validation or back testing involves looking at the output of a model from a point in time T and seeing how it performs at future time periods $T + t$. Out-of-sample validation helps to monitor the ongoing appropriateness of the model in light of what actually happens. In the case when markets experience sudden shocks (e.g., stock market crashes or economic crises), this type of validation may also help give confidence that tail calibrations are adequate and that the probability of severe events being within the ESG is adequate.

An example of how out-of-sample validation might look is shown in Figure 8.2. In this diagram, the so-called funnel of doubt from a particular model of municipal bonds estimated at month-end June 2012 is shown and compared with quarterly returns on BBB-rated municipal bonds post June 2012. The look forward period includes the Detroit default, which led to severe drawdowns in municipal bonds, as can be seen from the black line in Figure 8.2. The useful information contained in this plot is that the probability of the third-quarter 2013 loss with this particular model implies a probability of less than 1:200. In this particular case, it was determined that the deficiency was the result of a misspecification of the model, in that it contained no credit component. In other cases, such a “miss” may not be so directly attributable, and may be due to misspecification (i.e., the model is incapable of capturing such events), misparameterization (i.e., the model is capable of capturing such events but must be reparameterized to do so) or attributable to effects within the market or data being considered (e.g., special circumstances such as central bank intervention). In any case, the materiality of such results and what action if any is required need to be considered.

Figure 8.2 shows out-of-sample validation of BBB municipal bond returns between June 2012 and March 2014. The model shown uses only the tax rate and Treasury rates as explanatory variables of municipal bond yields, and no credit component is used.

Figure 8.2
OUT-OF-SAMPLE VALIDATION OF BBB MUNICIPAL BOND RETURNS



Sources: Bloomberg; Conning, Inc.

The combination of these validation techniques is used to give a fuller picture of the appropriateness of the model output and ultimately the robustness of the ESG as a whole. This and the aforementioned acceptance criteria are typically used together to continually monitor the ESG performance.

8.2.2 WHAT SHOULD BE VALIDATED?

So the task will be manageable, validation should take account of a broad set of simulated variables with particular focus on the areas of greatest materiality. Validators should also not restrict themselves to validating only the primary simulated variable (e.g., yield), but also any derived quantities (e.g., returns and excess returns) in order to get a fuller picture of the quality of the ESG. Also, it might be important to consider output such as yield curves to validate the dynamic properties of the ESG. For instance, does the ESG produce realistic yield curve shapes and inverted yield curves? For credit, special consideration needs to be given to the specific features of the credit market, such as transition and default probabilities, recovery rates and the adequacy of tail calibrations for spread distributions.

Correlation and interdependency between variables is another important area to validate. However, meaningful validation is not as simple as just measuring correlation from the available history and comparing it with the simulated value. Instead, decisions whether to accept a particular ESG calibration should take into account several considerations, such as the following:

- Length and quality of the available historical data
- Significance of values measured from historical data through comparison with other economies that have longer or better-quality data
- Statistical error or standard error on the measured historical correlation value
- Comparison of measured values with the average measured value of the ESG
- Range of correlation values produced by the ESG and comparison of these with the historical values
- Potential for overestimating the diversification benefits of two asset classes
- Extent to which it is likely that significant improvements can be made within the modeling framework

This sort of validation requires the validator to analyze the ESG output, compare it with historical data, understand the limitations of both the market data and the ESG models being used, and make an informed decision around the materiality of any differences.

8.3 VALIDATION OF MARKET-CONSISTENT ESGS

Market-consistent scenarios are typically used for the valuation of liabilities with embedded options and guarantees. The calibration of a model for these market-consistent embedded value (MCEV) calculations involves fitting the model at time zero to the market prices or implied volatilities of traded instruments such as swaps, swaptions, equity options and inflation derivatives. Unlike the validation of real-world scenarios, where there are a great many quantitative and qualitative possibilities for determining the quality of a scenario set, the validation techniques for market-consistent scenarios are fairly well defined. This is because the aims of an MCEV calibration are quite clearly defined:

- To fit the model to the market prices of instruments at time zero
- To show that the resulting scenarios are risk neutral

Thus, the model must only be validated for market consistency and risk neutrality.

The validation of each of these elements will now be discussed in more detail.

8.3.1 VALIDATING MARKET CONSISTENCY

Validating market consistency involves first computing the Monte Carlo prices implied by the model parameters. To do this, one runs a simulation of N stochastic paths or trials and evaluates:

$$P(t = 0) = E[D \cdot X]$$

where D is the discount factor and X is the expected payoff of the derivative or instrument in question. Practically this involves computing:

$$P(t = 0) = \frac{\sum_{i=1}^N D_i \cdot X_i}{N}$$

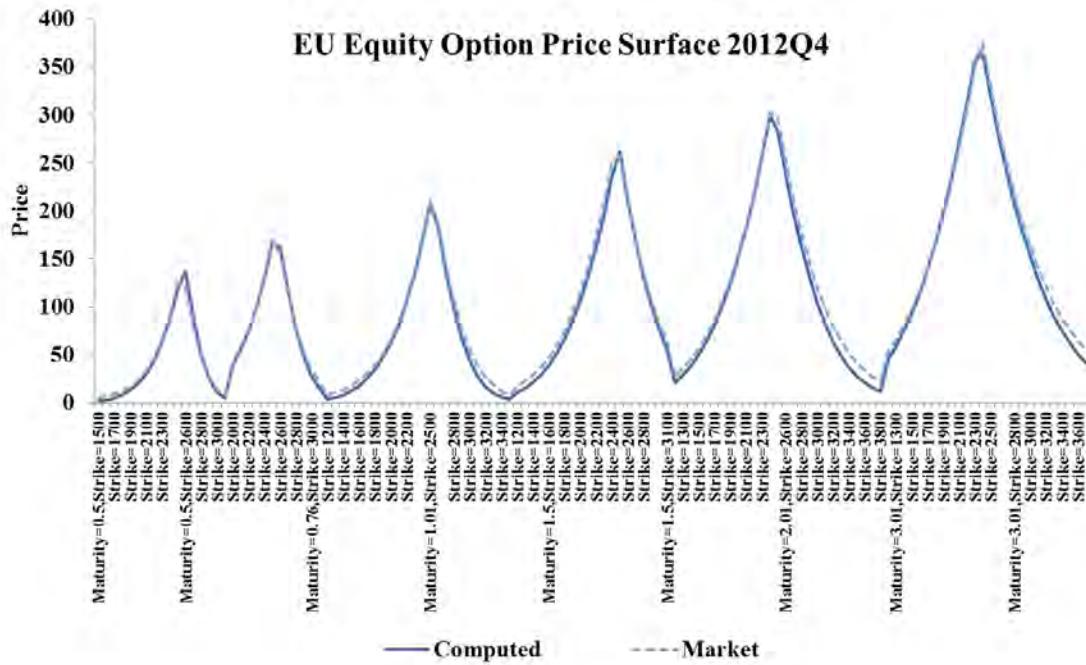
and comparing the value of this Monte Carlo price with the market price. An example of how this looks is shown in Figure 8.3, where the Monte Carlo prices of equity put options with different expiry dates and strikes for a particular model are compared with the market prices at a particular date. There is no market standard for how one assesses the quality of fit; however, suitable validation might include these tasks:

- Visual inspection and qualitative judgment on the adequacy of the fit
- Definition of particular acceptance criteria for the fit (e.g., particular instruments should be fit with a relative error of less than 10%)
- Use of a goodness of fit statistic

It is also important to point out that validation and acceptance should concentrate on the areas of the fit where the highest risk exposure exists. There is little to be gained by rejecting a model calibration if the fit is poor for options that are unrepresentative of the liabilities held. Hence the use of acceptance criteria and goodness of fit statistics should only ever be used as information that may act as a trigger for further investigation.

Figure 8.3 shows market-consistent calibration of a stochastic volatility model for equities (computed) to the market value of traded Eurostoxx 50 put option (market) at year-end 2012.

Figure 8.3
EQUITY OPTION PRICE SURFACE (COMPUTED VERSUS MARKET)



Sources: Bloomberg; Conning, Inc.

8.3.2 VALIDATING RISK NEUTRALITY: THE MARTINGALE TEST

Validating that the scenarios produced are in fact risk neutral is the second important task. This means the expected return on a particular asset class is the same as the expected return on the risk-free asset. A standard test known as a martingale or $1 = 1$ test is used to assess the extent to which this is true. In a $1 = 1$ test, the market value of a given self-financing portfolio (i.e., a portfolio that has no inflows or outflows of cash during the simulation) is simulated within the ESG across N stochastic paths. The test statistic at a given time t , m_t , is given by the following formula:

$$m_t = \frac{\sum_{i=1}^N D_{t,i} (1 + R_{t,i}^{\text{cum}})}{N}$$

where $D_{t,i}$ is the discount factor (implied by the risk-free asset) on path i at time t , and $R_{t,i}^{\text{cum}}$ is the cumulative return on the asset or portfolio between $t = 0$ and t . If the scenarios are risk neutral, and by extension the model arbitrage free, $m_t = 1$ for all times t . In reality, however, sampling error exists for the sizes of simulation typically used for this purpose, and $m_t = 1$ only as N approaches ∞ .¹ Therefore, it is important to base the acceptance criteria on a statistical measure rather than the absolute value of m_t . A common way to approach this is to evaluate the standard error of m_t .²

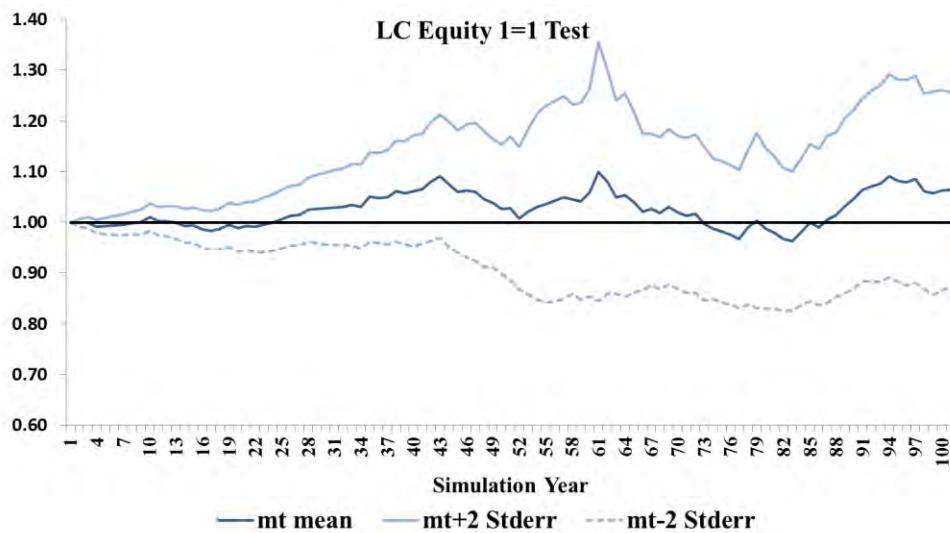
$$\text{SE}_{m_t} = \frac{\sigma_{m_t}^2}{\sqrt{N}}$$

¹ Although other sources of error such as discretization error and the pseudo-randomness of the random number generator may cause m_t to deviate from unity.

² Formula is shown in the absence of variance reduction techniques such as antithetic draws.

and ensure that the value is consistent with 1 within X standard errors for all times t . X is often chosen to take a value of 2—i.e., that $(m_t + 2 \text{SE}_{m_t}) \geq 1$ and $(m_t - 2 \text{SE}_{m_t}) \leq 1$ —as an approximation to the 95% confidence interval. An example of how this might look graphically is shown in Figure 8.4. Neither of the lines representing the two-standard-error interval crosses the black line representing $m_t = 1$, so the calibration would be said to pass this particular validation test. Figure 8.4 shows a $1 = 1$ test result on a calibration to a large-cap equity option. The upper and lower bounds represent the two-standard-error interval.

Figure 8.4
1 = 1 TEST RESULT

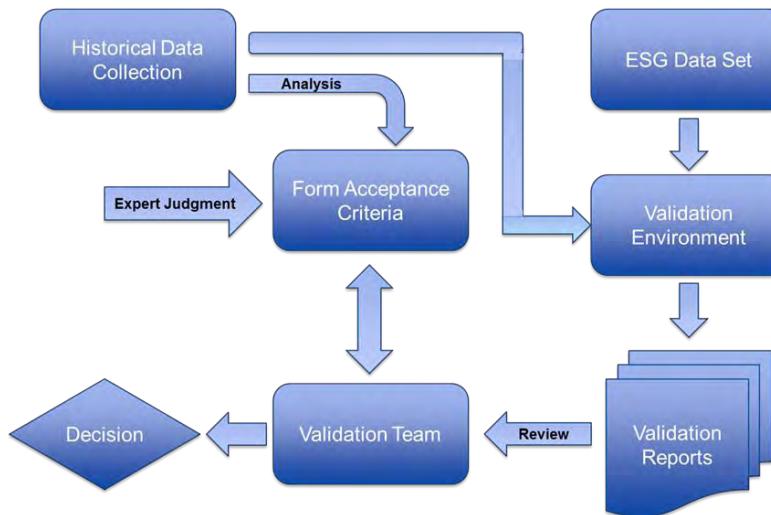


Source: Conning Inc.

8.4 VALIDATION: FINAL THOUGHTS

Figure 8.5 shows the steps involved in a typical validation process. This validation is vital to ensure that models are appropriate to the application under consideration for both real-world and market-consistent ESGs. However, the validation of a real-world ESG is different from that of a market-consistent ESG. In the validation of a real-world ESG, a great deal of scope exists with regard to the choice of variables to validate, the modes of validation and the forming of acceptance criteria, which may include expert judgment. For a market-consistent ESG, the aim of the validation and the mode of analysis are much more clearly defined, and market standards exist with respect to this. Nevertheless, both real-world and market-consistent validation require, to some extent, the availability and selection of appropriate data, and systems in place to maintain and use that data in an efficient and accurate manner.

Figure 8.5
MAIN STEPS IN AN IDEALIZED VALIDATION PROCESS



Source: Conning Inc.

In practice, the most effective validation is one that combines multiple modes of analysis. For real-world ESGs, point-in-time, in-sample and out-of-sample validation techniques, combined with well-defined acceptance criteria, should be used together to determine not just what is wrong with a model, but also what is right with the model. Such a holistic approach gives the validator a global view of the performance and robustness of the ESG. Finally, the importance of human intervention in the validation process should not be underestimated. A validator with experience of the structure, limitations and usage of the model who is able to combine both quantitative and qualitative judgments is likely to be far more effective than reliance on a particular set of rigid criteria or goodness of fit statistics to determine the robustness of the model.

8.5 CHAPTER SUMMARY

1. Validation is required for both real-world and market-consistent calibrations, and parameterizations and validation procedures differ between the two.
2. An automated validation system is preferable to manual validation.
3. The validation process should start with acceptance criteria and then move on to the validation stage.
4. Validation of real-world scenarios requires data and expert judgment.
5. Validation of real-world ESGs against market data falls into three distinct categories: point-in-time validation, ex-post or in-sample validation, and ex-ante or out-of-sample validation.
6. Market-consistent scenarios need only be validated for market consistency and risk neutrality.
7. Market consistency involves computing Monte Carlo prices and comparing these prices with market prices.
8. Risk neutrality is validated using a martingale test.
9. Observation by a suitably trained and expert user is vital for both real-world and market-consistent validation.

8.6 CHAPTER REFERENCES

Christodoulakis, George A., and Stephen Satchell. 2007. *The Analytics of Risk Model Validation*. Burlington, MA: Academic.

Scandizzo, S. 2016. *The Validation of Risk Models: A Handbook for Practitioners*. London: Palgrave Macmillan.

Chapter 9: Arbitrage-Free Modeling Considerations

Arbitrage is the opportunity to make a riskless profit by exploiting price differences of identical or similar financial instruments, on different markets, or in different forms. The existence of arbitrage is sometimes referred to as a “free lunch.” One of the key concepts in modern finance is the notion of the absence of arbitrage. The operational manifestation of this concept is an indispensable tool in the classical construction of ESG models.

The condition of designing an arbitrage-free financial model is both reasonable and desirable so long as we believe that markets are efficient. Models that are not arbitrage-free can be useful, but they need to be interpreted carefully. For example, Redington’s classical immunization theory remains a useful part of the actuarial tool kit even though one can easily construct a free lunch from barbell strategies.

Section 9.1 of this chapter provides a definition of arbitrage-free financial models, referencing cash flow and trading strategy mechanisms to assure that the models are consistent. The theory of an arbitrage-free model requires that all assets in the model have well-defined cash flow structures. Pricing can then be determined by first identifying a trading strategy that generates the same cash flows as the cash flows that need to be priced. The price of this cash flow stream is equal to the known prices of the securities that replicate that cash flow. The model is said to be arbitrage-free if there do not exist any self-financing trading strategies that produce a risk-free gain. The section presents some basic mathematics to describe the pricing of cash flow streams in the case of a zero-coupon bond and a non-dividend-paying stock.

Section 9.2 describes, in mathematical notation and words, how to define the results of trading strategies and account for all cash flows as trading positions change. It introduces the concept of hedging a cash flow stream by a trading strategy. The cash flow stream is hedged if it can be exactly replicated by dynamically trading in the securities market. In an arbitrage-free model, prices of assets are discounted expectations across all future events. Arbitrage-free asset prices therefore embed all possible future behavior in today’s prices. This forward-looking aspect of arbitrage-free models is one reason that prices cannot be prescribed directly.

Section 9.3 discusses some of the considerations of whether an arbitrage-free framework is necessary and even still relevant. The dynamics that can be used to drive risk are constrained; the dynamics that can be applied may not lead to a reasonable pricing formula in today’s experience. One area that is getting a lot of attention is the need for more robust interest rate models that accommodate the ultra-low interest rates we are currently experiencing. If ESG models are applied in real-world modeling contexts, then these models should be able to capture return behavior and market dynamics but do not necessarily have to be technically arbitrage-free. But if the scenarios are not arbitrage-free, then one cannot consistently price derivatives.

Section 9.4 expands on the martingale test described in the previous chapter. Martingale tests can be used to confirm that the scenarios generated by an ESG simulation are arbitrage-free. The application of a martingale test is predicated on the assumption that risk-neutral output is available for the model to be tested. The simplest application of the martingale test is for assets that do not generate cash flows. If the scenarios are risk-neutral, the time zero price of the asset should equal the average of the present value of the ratios of the risk-neutral prices and the values of the asset. When the risk-neutral output for a dividend-paying asset is to be tested, the martingale test can be based on the expected value of the discounted dividend stream over the horizon plus the discounted value of the asset at the horizon.

9.1 ARBITRAGE-FREE FINANCIAL MODELS

The importance of the theory of arbitrage-free financial models for ESG work is that it provides a mathematical framework that can be used to verify an ESG’s consistency. The cost that must be paid for this framework is hard mathematics. The mechanism through which consistency is assured depends on the relationship between the cash flows of securities that are modeled and traded, and the prices of these securities. In practice, it is usually the cash flows for a security that are readily described, and the theory focuses on how the price of these cash flows must behave. Furthermore, the restrictions that the absence of arbitrage imposes on the price of these cash flows are useful in developing models for the prices of securities that populate the ESG.

The theory of an arbitrage-free model (such as would be found in an ESG) requires that all assets in the model have well-defined cash flow structures. For example, if an equity index is to be included in the model, then its ex-dividend price process and dividend cash flow structure must be specified. Two important concepts are then introduced to the model: a modeled set of trading strategies, and self-financing trading strategies. The purpose of modeled trading strategies within the ESG is to establish a mechanism of buying and selling securities. Self-financing trading strategies involve the construction of specific trading strategies that do not require the injection of additional capital during their execution and for which all cash flows generated by the trading strategy are accounted for. The difference between the initial cost of the trading strategy and its value at a given point in time is referred to as the gains process; it represents the gain or loss created by the trading strategy from its inception to that point in time. An arbitrage is then defined as an explicit trading strategy involving the simultaneous buying and selling of securities, possibly quite complicated, leading to either of the following outcomes:

- The trader would receive a positive cash flow today and no further liabilities during the rest of the trade horizon.
- The trader could enter into the trade at zero net cost and receive only nonnegative cash flows during the rest of the trade horizon, with a positive cash flow occurring with probability greater than zero.

The model is said to be arbitrage-free if there do not exist any self-financing trading strategies that produce an arbitrage. While the financial notions associated with these ideas are clear enough, these concepts are not useful for ESG purposes until they are given a mathematical interpretation that can be imposed on the model-building process. At this point, the theory becomes very technical.

If a model is arbitrage-free, then one can assign a price to a cash flow stream according to the principle that the price of a cash flow stream is equal to the price of any self-financing trading strategy that generates the same cash flow stream. In an arbitrage-free model, securities producing identical cash flow streams are essentially interchangeable from a pricing standpoint. The logic is simple; two identical cash flow streams ought to have the same price. If they do not have the same price, then one could create an arbitrage by selling the more expensive one and covering the position by buying the cheaper one. Implicit in this pricing notion is that there exists a self-financing trading strategy that generates the cash flow stream, and that we know the prices of the securities that need to be traded in order to generate the cash flow stream.

The significance of the first point is that we can obtain the cash flow stream we are interested in by trading in the available securities. The significance of the second point is that we are able to associate a price to the cash flow stream based on knowing the prices of the securities that its cash flows are derived from. In summary, we are able to price derivative securities based on determining a trading strategy that generates the cash flows we want to price and then using the known prices of the traded securities to infer the price of the cash flow stream.

Under reasonable mathematical conditions it is shown that a model is arbitrage-free if and only if there exists a change of measure under which the discounted gains process for all traded assets, for any trading strategy, is a martingale (i.e., a stochastic process in which the conditional expectation of the next value, given the current and preceding values, is the current value). Martingale tests were discussed in Section 8.3.2 and are further discussed later in this chapter. This result may be stated using a range of terminology, but the way it is most commonly thought of is in the language of “risk-neutral expectation.”

It is possible to understand the general operational details without getting deeply into mathematical theory. We will therefore sidestep some of the theoretical justifications for the results we are discussing. However, the concept of arbitrage-free models is a precise and technical one, and we require some mathematical notation in order to proceed. Also, the difficulty level of the material must increase accordingly. Chapter 10 gets into more detail about risk-neutral scenarios, including a more precise definition of risk neutrality and some of the mathematical framework around finding a price and replicating portfolio in an arbitrage-free context.

Let suppose that there is defined a continuously compounded instantaneous rate of interest, which we will denote by r_t . Actuaries are very familiar with this idea from their early studies of compound interest theory, where this quantity is classically referred to as the “force of interest.” It is usually assumed to be constant and is denoted by the small Greek letter δ . If one invests a unit amount (e.g., a dollar) for t time units at a continuously compounded instantaneous rate of interest δ , then the value of the investment at time t is $\exp(\delta_t)$. If one invests a unit amount for

t time units at a stochastically fluctuating continuously compounded instantaneous rate of interest r_u , then the value of the investment at time t is $\exp(\int_0^t r_u du)$. We may think of this investment as a bank account, and we may define a process:

$$B(t) \stackrel{\text{def}}{=} \exp\left(\int_0^t r_u du\right)$$

which represents the accumulated value of the initial deposit of one unit into the bank account over a period of t years. This bank account process will serve as our means of discounting.

The central question we are faced with is, How can a cash flow stream be priced in an arbitrage-free manner? The general answer is that if the price of the cash flow stream is to be consistent with the absence of arbitrage, then the price must be given as the expected discounted future cash flows, where the expectation is taken with respect to a risk-neutral measure. The precise formulation of this result has come to be known as the First Fundamental Theorem of Asset Pricing. For example, in the case of a zero-coupon bond maturing for a unit at time T , the cash flow stream consists of a single lump sum payment at time T , and its price must be of the following form:

$$P(0, T) = E^*\left[\frac{1}{B_T}\right] \equiv E^*\left[\exp\left(-\int_0^T r_u du\right)\right]$$

where $P(0, T)$ denotes the price at time zero of a unit amount that is paid with certainty at time T . In the case of a non-dividend-paying stock with price process $\{S_t\}$, its price must satisfy the following formula:

$$S_0 = E^*\left[\frac{1}{B_T} S_T\right] \equiv E^*\left[\exp\left(-\int_0^T r_u du\right) S_T\right].$$

When the stock pays dividends, the corresponding formula is more complicated and must include the expected value of the discounted dividend stream.

9.2 APPLYING A PRACTICAL APPLICATION

To put the theory on a firm foundation, we must be able to define the results of trading strategies and have a way to account for all cash flows as trading positions change. This requirement naturally leads to integrals of the trading positions $\{\vartheta_u | u \in [0, T]\}$ weighted by the change in the value of the position dS or the cash flow paid by the security dD , leading to evolving gains from trade of the form

$$\int_0^t \vartheta_u dS_u + \int_0^t \vartheta_u dD_u.$$

The cash flow stream $\{C_u | u \in [0, T]\}$ is said to be hedged if there exists a trading strategy $\{(\phi_u, \vartheta_u) | u \in [0, T]\}$ satisfying the following expression:

$$\phi_t B_t + \vartheta_t S_t = \phi_0 B_0 + \vartheta_0 S_0 + \int_0^t \phi_u dB_u + \int_0^t \vartheta_u dS_u + \int_0^t \vartheta_u dD_u - \int_0^t dC_u$$

for all $t \in [0, T]$ and for which

$$\phi_T B_T + \vartheta_T S_T = 0.$$

In words, the meaning of this equation is that the cash flow stream is hedged if it can be exactly replicated by dynamically trading in the securities market. The secondary condition says that all capital has been used up when the cash flow stream expires.

If a cash flow stream is hedged by the trading strategy $\{(\phi_u, \vartheta_u) | u \in [0, T]\}$, then, in an arbitrage-free market, it is natural to define the price of the cash flow stream as follows:

$$\phi_0 B_0 + \vartheta_0 S_0$$

which is the cost of the trading strategy that generates the cash flow stream. The way in which the theory assigns a price to a cash flow stream is to take the price to be the cost of the trading strategy that generated these cash flows.

What would make this definition operationally useful is for there to be a convenient way to compute the previous expression. Using stochastic calculus, it may be shown that the previous expression can be computed from the following formula:

$$\text{Price}(\{C_u\}) = E^* \left[\int_0^T \frac{1}{B_u} dC_u \right]$$

The value of this formula is that the price of the cash flow stream is now expressed in terms of the discounted cash flows themselves, rather than in terms of the unknown quantity in the prior formula. We may say that the price of the cash flow stream must satisfy this formula if it is to be consistent with the absence of arbitrage.¹

A related formula is often used based on the concept of state price density. The state price density is a positive process that may be used to price cash flow streams according to this formula:

$$\text{Price}(\{C_u\}) = E \left[\int_0^T \rho_u dC_u \right].$$

The expectation is taken under the real-world measure, which means the probability distributions that define all security dynamics are consistent with the way we actually experience them. While the second expression appears simpler, because there is no risk-neutral measure, in practice the first formula is usually applied because the risk-neutral change of measure can be obtained by a change in the model parameters only.

The equivalence between the change of measure to risk neutral and the appropriate change of model parameters is what permits actuaries without extensive knowledge of stochastic calculus to still work effectively with a suite of ESG models.

The other important aspect of the theory relates to the spanning of cash flows. Not all cash flows can be created through dynamic trading. For example, one cannot replicate the cash flows from catastrophe risk bonds by dynamically trading in equity indices and government bonds, no matter how intricate the trading strategies may be. This is because the behavior of the cat risk bond depends on variables not present in the equity and government bond market.

The technical terminology that the reader will encounter in delving further into this theory is market completeness. The issue of completeness is relevant to ESG applications that require derivatives pricing and hedging. We will note some of the high-level issues that relate to the more sophisticated equity models in Chapter 13.

The formula for asset prices in terms of discounted expectations we have discussed imposes strong restrictions on ESG models. An actuary cannot arbitrarily invent and impose a realistic asset price dynamic and still have an arbitrage-free model. In an arbitrage-free model, prices of assets are discounted expectations across all future events, and arbitrage-free asset prices therefore embed all possible future behavior in today's prices. This forward-looking aspect of arbitrage-free models is one reason that prices cannot be prescribed directly. Indeed, prices must embed the range of outcomes that the model is allowing for in the future, and the price we observe today is an expectation of these outcomes.

9.3 IS THE ARBITRAGE-FREE FRAMEWORK NECESSARY?

The application of the arbitrage-free modeling approach comes with benefits as well as restrictions. The restrictions can be inconvenient: One is constrained in the dynamics that can be used to drive risk, and the dynamics that can be

¹ The integral is stochastic and can be developed under different frameworks. The rules for the manipulation of the integrals, "the stochastic calculus," and the conditions under which the integral exist vary with the type of integral used. In all cases, the financial content is the same in that the integral serves to aggregate the discounted cash flows and there is no loss of fidelity in thinking of the integral in calculus terms.

applied may not lead to tractable pricing formulas. The events of the last decade have placed incredible pressure on ESG models, and one area that is getting a lot of attention is the need for more robust interest rate models that accommodate the ultra-low interest rates we are currently experiencing. Many interesting dynamics could be useful in better capturing ultra-low interest rates, but they are not easily brought into the arbitrage-free framework. This type of consideration is one argument for moving away from the arbitrage-free requirement.

If ESG models are applied in real-world modeling contexts, then these models should be able to capture return behavior and market dynamics but do not necessarily have to be technically arbitrage-free. Broadly speaking, one could construct a defensible ESG that is not technically arbitrage-free, so long as the model produces returns that relate to actual market behavior and respects general relationships between financial variables.

However, arbitrage-free models are a necessary requirement when one wants to apply an ESG to risk-neutral pricing problems. If the scenarios are not arbitrage-free, then one cannot consistently price derivatives. It would be difficult to justify the use of models that were not arbitrage-free in risk-neutral pricing and hedging applications.

9.4 MARTINGALE TESTS

When risk-neutral ESG simulation output is presented, a standard test that is applied to confirm that the scenarios are risk-neutral is the martingale test. The martingale test is based on the fact that in a risk-neutral framework, there is an explicit relationship between the price of a cash flow stream and its expected discounted future cash flows. This relationship can be exploited in a variety of ways to derive tests in simulated output that constitute a necessary condition for the scenarios to be risk-neutral.

The application of a martingale test is predicated on the assumption that risk-neutral output is available for the model to be tested. Under some circumstances, useful econometric models are employed in ESGs for asset returns or other variables that may not have a risk-neutral equivalent. The martingale test will not apply to such models.

The martingale test can be applied to a single asset such as a stock index or a bond. The test can also be applied to the returns on an index constructed from a collection of traded assets. The specific mechanics of the test varies depending on the type of asset or return index the test is applied to.

The simplest application of the martingale test is for assets that do not generate cash flows, standard examples being non-dividend-paying stocks or zero-coupon bonds. Risk-neutral prices are generated for the assets, and the risk-neutral short rate of interest also is generated. A time horizon is selected. The risk-neutral short rates are used to compute the value of the bank account process for the time horizon, and the risk-neutral price of the asset is divided by the bank account value. If the scenarios are risk-neutral, the average of these ratios should be equal to the time zero price of the asset.²

When an asset that generates cash flows is tested, the same general procedure applies, with some refinements necessary to deal with the more complicated cash flow structure. When the risk-neutral output for a dividend-paying asset is to be tested, the martingale test can be based on the expected value of the discounted dividend stream over the horizon plus the discounted value of the asset at the horizon. The average of these values should be equal to the time zero price of the asset. Alternatively, the test can be based on the expectation of the discounted value of the adjusted asset position at the horizon where the adjusted asset position reflects the dividend payments under the assumption that these are reinvested in the asset over the horizon.

Martingale tests are also frequently applied to total return series. It can be shown, using the ideas from this chapter, that if a total return index is based on a portfolio of traded assets, then the expected discounted value of that index at a given horizon will also be equal to the time zero value of the index. The index inherits this property from the constituent assets that make up the index. The fact that this test is being performed on a total return index means that cash flows from assets in the index are being reinvested into the index.

² Since there is random sampling error, the average will never be exactly equal to the time zero price of the asset, but for a relatively large sample size, the average will be close, provided that the scenarios are risk-neutral. If this average is not close to the time zero price of the asset, then the ESG fails the martingale test, and this indicates that the scenarios are not risk-neutral. This might indicate a parameterization error, the use of a model that does not have a risk-neutral risk version, or some other problem with the ESG.

9.5 CHAPTER SUMMARY

1. The concept of an arbitrage-free financial model is both reasonable and desirable, so long as we believe that markets are fairly efficient. Models that are not arbitrage-free—the classical actuarial example being Redington's theory of immunization—can still be very useful but need to be interpreted carefully.
2. The practitioner view on whether ESG models should be arbitrage-free has evolved over time. Prior to the financial crisis, most practitioners would generally agree that ESG models should be arbitrage-free. The peculiar economic nature of the past many years has practitioners clamoring for “flexibility” in their interest rate models, and one can build all sorts of interesting real-world interest rate dynamics, so long as one is unfettered by the need for the absence of arbitrage. Some model discipline needs to be imposed in order to avoid unreasonable risk-return trade-offs, but there is room for debate on both sides of the issue.
3. The technical meaning of arbitrage-free is expressed in terms of trading strategies. The mathematical formulation of trading strategies and the handling of the cash flows from traded assets are what gives rise to the fundamental theorem of asset pricing. It is at this point that stochastic calculus and martingale theory enters ESG theory.
4. Using stochastic calculus, one can express the price of a cash flow stream in terms of the discounted cash flows themselves. This requires a risk-neutral environment.
5. An alternative approach is based on the concept of state price density, which operates under the real-world measure, which means that the probability distributions that define all security dynamics are consistent with the way we actually experience them. In practice, the risk-neutral measure is more commonly applied, because the formula is simpler to use.
6. The use of trading strategies and the concept of self-financing lead to the characterization of asset prices as discounted expectations (i.e., risk-neutral expectations). The precise nature of the discounted expectation depends on the nature of the cash flows associated with the asset to be priced (non-dividend-paying, dividend-paying, defaultable, etc.).
7. Formulas for asset prices in terms of discounted expectations can be very useful for model building. A standard example is the expression of zero-coupon bond prices in terms of the expectation of the integral of the “short rate.”
8. The formula for asset prices in terms of discounted expectations also imposes strong restrictions on ESG models, so that an actuary cannot arbitrarily invent and impose a realistic asset price dynamic and still have an arbitrage-free model. In an arbitrage-free model, prices of assets are discounted expectations across all future events, and arbitrage-free asset prices therefore embed all possible future behavior in today's prices.
9. The formula for asset prices in terms of discounted expectations leads to a martingale condition for asset prices. This martingale condition can be used as a test of ESG scenarios as a necessary condition for the scenarios to be risk-neutral. This test is widely applied by regulators and is generally referred to as “the martingale test.”
10. Setting up a martingale test requires care to properly account for the nature of the cash flows for the asset being tested. The construction of the test differs for treasury bonds, equity returns and corporate bonds.

9.6 CHAPTER REFERENCES

The material in this chapter goes back to the pioneering work of Fischer Black, Myron Scholes, Robert Merton, Steve Ross, and others. The formalization of these ideas in continuous time goes back to the work of J. Harrison and D. Kreps and was taken up by many others. Foundational work may be found in the following sources:

Harrison, J. Michael, and David M. Kreps. 1979. Martingales and Arbitrage in Multiperiod Securities Markets. *Journal of Economic Theory* 20: 381–408.

Harrison, J. Michael, and Stanley R. Pliska. 1981. Martingales and Stochastic Integrals in the Theory of Continuous Trading. *Stochastic Processes and their Applications* 11: 215–260.

Redington, F. M. 1952. Review of the Principles of Life Office Valuations. *Journal of the Institute of Actuaries* 78: 286–340.

Many texts can be used to learn about this theory. There is no escaping some level of mathematical sophistication for continuous-time models.

Björk, Tomas. 2009. *Arbitrage Theory in Continuous Time*. 3rd ed. New York: Oxford University Press.

Duffie, Darrell. 2001. *Dynamic Asset Pricing Theory*. 3rd ed. Princeton, NJ: Princeton University Press.

Chapter 10: The Role of Risk-Neutral Scenarios

As we saw in Chapter 9, the concept of arbitrage-free models provides a set of useful restrictions on the securities in an ESG. An essential result that emerges from this framework is that the price of a traded security at any point in the simulation must be equal to the expected value of its future discounted gains process under a risk-neutral measure. Risk-neutral theory provides a mathematical framework to find the price and replicating portfolio in an arbitrage-free context. For some insurance-pricing applications, the analysis ends here. In other applications, the prices and replicating portfolios are then applied in a real-world environment.

This chapter is laid out as follows:

Section 10.1 describes what risk-neutral means, including the expansion on the concept of a martingale and a martingale test that was introduced in prior chapters. Risk-neutral does not depend on developing qualitative behavior of financial variables that conform to historical statistics. Instead, risk-neutral focuses on the ability to fit financial variables to today's market prices. For an asset that pays no dividends, risk-neutral requires calibration of parameters so that the expected return (over a short period of time) is equal to the risk-free interest rate. When more general assets are considered, risk-neutral means that the parameters are calibrated so that the expected discounted gains process for that variable is the same for all time horizons. A risk-neutral scenario set can be interpreted as pricing scenarios because it permits a series of cash flows, under varying economic scenarios, to be discounted at a risk-free rate to produce prices of the cash flows. The relationship between the price of a cash flow stream and the expectation of its discounted future cash flows for a risk-neutral scenario set provides a framework for the direct auditing of model output through a martingale test.

Section 10.2 introduces some examples of applications of a martingale test in practice. Whether a martingale test is for a non-dividend-paying asset, a dividend-paying asset or a total return index, the test output appears similar. Martingale tests are based on averages of random quantities. Therefore, we expect there to be random fluctuations in the test output. This section provides comparative examples of typical results of martingale tests of a moderate-volatility asset and a high-volatility asset, as well as an example of an inconclusive martingale test of a high-volatility asset.

Section 10.3 develops why we need risk-neutral scenarios and provides some examples of what risk-neutral scenarios are used for. ESG models under risk-neutral scenarios permit us to evaluate the current price of an asset as the expected discounted value of the future cash flows associated with that asset. Under risk-neutral scenarios, this expectation can be evaluated in closed form for a wide range of assets. It is also useful to be able to parameterize a model to exhibit qualitative behavior that is consistent with market norms. However, in many important insurance applications, the risk-neutral expectation that gives the price of an insurance liability cannot be readily evaluated in closed form, but we can compute the expectation for each scenario across all scenarios and average them to get the expected value that gives the price. This procedure is commonly referred to as Monte Carlo simulation. It is also useful to be able to parameterize a model to exhibit qualitative behavior that is consistent with market norms.

Section 10.4 discusses considerations of when risk neutral scenarios should be used, and when real-world scenarios should be used. The general distinction is that risk-neutral scenarios are used for pricing and real-world scenarios are used to assess risk. Banks that trade derivatives are primarily interested in the price of the derivatives they are buying and selling, and risk-neutral scenarios are often their primary focus. Risk-neutral scenarios are typically used when one must price cash flows for which closed formulas are not available, and the risk-neutral scenarios are applied to numerically compute prices by Monte Carlo simulation. Real-world scenarios are applicable for measuring returns, volatilities and other risk exposures. Risk-neutral scenarios are applicable for pricing cash flows. Real-world scenarios are calibrated to historical benchmarks or clients' own views. Risk-neutral scenarios are calibrated to pricing data such as swaptions and index options.

Section 10.5 discusses examples of when risk-neutral scenarios and real-world scenarios are used, together with some of the practical considerations with inner loops. Risk-neutral scenarios will be used in conjunction with real-world scenarios when the real-world risk measurement exercise requires prices for liabilities or assets that do not have closed-form formulas. Such applications typically involve stochastic on stochastic (i.e., inner loops) where the

state of the world or node is simulated under the real-world measure and cash flows are priced in that state of the world using risk-neutral scenarios projected forward from the current node of the simulation. One purely financial example that occurs in practice is managing portfolios of mortgage-backed securities (MBS).

10.1 WHAT DOES “RISK-NEUTRAL” MEAN?

Our primary focus in this section is to discuss the technical meaning of the term *risk-neutral* and explain how this leads to the typical interpretation of risk-neutral scenarios as pricing scenarios. Each of the financial variables that are modeled in an ESG will have a specified stochastic process that drives the movement of that variable. The parameters of that stochastic process can be chosen or calibrated to produce certain qualitative behavior. In real-world risk management, calibration is done to obtain qualitative behavior that matches selected targets. The nature of the qualitative behavior in a risk-neutral context is more subtle. Risk-neutral scenarios are more difficult to interpret, because they are not intended to conform to historical distributions. Their primary criterion is the ability to fit today's market prices.

Consider the case of an asset that pays no dividends. Profit or loss from trading this asset will be determined solely by changes in the price of the asset. For this type of asset, risk-neutral means that the parameters are calibrated so that the expected return from this asset over a short period of time is equal to the risk-free interest rate.

When more general assets are considered, risk-neutral means the parameters are calibrated so that the expected discounted gains process for that variable is the same for all time horizons. It is this last notion that can be more precisely expressed in the language of martingales. The use of martingales is not a mere convenience, because technical results from martingale theory are needed to understand how to set the parameters to give risk-neutral behavior. Furthermore, understanding why the concept of risk-neutral is useful depends on arbitrage-free pricing theory, which was discussed in the previous chapter. We now further discuss these issues.

Using the notation from the previous chapter, if V_t represents the value of a portfolio generated by a self-financing trading strategy, then we have the following relationship:

$$\frac{V_0}{B_0} = \int_0^t \vartheta_u dM_u + \frac{V_t}{B_t}$$

$$M_t = \frac{S_t}{B_t} + \int_0^t \frac{1}{B_u} dM_u.$$

The first expression simply says that the discounted value of a portfolio generated by a self-financing trading strategy is equal to the initial investment plus gains from trade as captured by the integral term. If there are parameters for the process M_t in the second equation, so that

$$E \left[\int_0^t \vartheta_u dM_u \right] = 0$$

then taking expectations in and canceling the term involving M_t gives:

$$\frac{V_0}{B_0} = E \left[\frac{V_t}{B_t} \right].$$

This relation says we can compute the price of a cash flow stream or security by taking the expectation across all future outcomes of its discounted value, providing that risk-neutral parameters are used in evaluating the expectation. The martingale concept is related to the condition identified in the first expression.

A risk-neutral scenario is a collection of simulated values of variables, each of which is parameterized by risk-neutral parameters. A risk-neutral scenario set is a collection of many risk-neutral scenarios. Each of the constituent risk-neutral scenarios in a risk-neutral scenario set is interpreted as a particular simulated outcome for the economy. A risk-neutral scenario set is therefore a complete simulation of various outcomes for the whole economy.

A risk-neutral scenario set can be interpreted as pricing scenarios, because it permits a series of cash flows, under varying economic scenarios, to be discounted at a risk-free rate to produce prices of the cash flows.

A common application is to the pricing of a zero-coupon bond. If the zero-coupon bond matures for a unit amount at time T , then its price at time zero must be given by an expression of the following form:

$$E \left[\exp \left(- \int_0^T r_u du \right) \right]$$

where we have replaced $1/B_T$ with the exponential term in the expectation and used the fact that $V_T = 1$ for a zero-coupon bond maturing for a unit at time T . In practice, many models for bonds permit explicit evaluation of this equation or efficient numerical processes such as the solution of ordinary differential equations to compute this equation. Generally, ESGs seek to exploit closed-form solutions and efficient numerical procedures for pricing whenever possible.

We now briefly turn to the idea of the martingale test. As we have previously noted, financial quants use martingales as a fundamental tool, and it is not possible to explain the nuts and bolts of arbitrage-free markets and the martingale test without them. It is possible to explain the practical implications of the theory, though, and that remains our focus, so we will briefly discuss the idea of a martingale and then move to the practical aspects of martingale tests.

In discrete time, a martingale is a sequence of random variables for which, at any point in time, the conditional expectation of the next value in the sequence is equal to the present observed value, given knowledge of all prior observed values (i.e., the history of the sequence). In a sense, the future value is not predictable, and knowledge of all prior values adds no information to the conditional expectation of subsequent values. The continuous time version of a martingale is a direct generalization of this idea. The observed history of the martingale is referred to as the information structure or filtration.

Martingale tests arise in ESG practice because they are often required by regulators. Regulators cannot check all of the fine details of an ESG in order to ascertain its veracity, and they are therefore interested in auditing tests that can be conducted directly on model output. The relationship between the price of a cash flow stream and the expectation of its discounted future cash flows for a risk-neutral scenario set provides a framework for the direct auditing of model output. The implementation of the martingale test requires a careful assembly of the relevant risk-neutral output. As we noted in Section 9.4, the risk-neutral output required to perform the martingale test depends on the type of asset being tested.

Constructing a martingale asset test begins with the choice of a horizon $T > 0$ and the following expression:

$$V_0 = E \left[\int_0^T \frac{1}{B_u} dC_u + \frac{V_T}{B_T} \right]$$

which is valid for self-financing trading strategies and for which it is assumed that all traded assets have risk-neutral parameterizations so that the expectation is a risk-neutral expectation. A specific self-financing trading strategy is now selected, and this selection is made to isolate a particular asset and serves to define the nature of that martingale test. For example, if one selects the self-financing strategy of purchasing an asset with ex-dividend price process X_t and reinvesting the dividends into that asset, then the martingale test at horizon T takes the following form:

$$X_0 = E \left[\frac{1}{B_T} Q_T X_T \right]$$

where Q_T denotes the accumulated units of the asset that are held at time T as a result of the reinvestment of dividends and the original investment. To apply the martingale test, one must determine Q_T . In cases of practical interest, Q_T is usually a deterministic function of a dividend rate or payment schedule and the horizon T . To apply the martingale test, one must give a risk-neutral scenario set consisting of the values for X_T and B_T and then compute the average:

$$\frac{1}{M} \sum_{i=1}^M \frac{1}{B_T(i)} Q_T(i) X_T(i)$$

which should be close to X_0 . We denote the number of scenarios in the risk-neutral scenario set as M .

The martingale test is often set up as a $1 = 1$ test. This is nothing more than normalizing both sides of martingale test at horizon T by X_0 , resulting in this test expression:

$$1 = \frac{1}{M} \sum_{i=1}^M \frac{1}{B_T(i)} Q_T(i) \frac{X_T(i)}{X_0}$$

A total return index may be constructed from traded assets according to this simple weighted formula:

$$1 + R(t, t+s) = 1 + \sum_i \omega_i R^i(t, t+s)$$

where $R^i(t, t+s)$ is the total return on asset i over the investment period beginning at time t and ending at time $t+s$ (i.e., over the interval $[t, t+s]$) and ω_i is the portfolio weight of asset i . The corresponding total return index is now constructed recursively, starting with $I(0) = 1$ and proceeding as follows:

$$I(t+s) = I(t) (1 + R(t, t+s)).$$

It can now be shown, using the martingale property of each of the constituent assets, that

$$I(0) = E\left[\frac{1}{B_T} I(T)\right]$$

at horizon T . This expression is a generic but easy-to-apply martingale test for total return indices. To perform the martingale test, one must have available a risk-neutral scenario set of the total return series for the index. This is not difficult to compute, but it must be computed correctly, or the premise of the martingale test will be violated, and the preceding equation will not hold.

10.2 APPLICATIONS OF MARTINGALE TESTS IN PRACTICE

Whether a martingale test is for a non-dividend-paying asset, a dividend-paying asset or a total return index, the test output appears similar. In this section, we look at some sample martingale test output to show what one expects from the tests in cases where the test is both passed and failed.

The first thing that needs to be noted is that martingale tests are based on averages of random quantities. Therefore, we expect there to be random fluctuations in the test output. For more volatile asset prices or returns, such as equity returns, one may get martingale test results that are not perfect passes, even though the scenarios may be legitimate risk-neutral scenarios.

The examples provided are based on a non-dividend-paying asset with a simple price process of the Black-Scholes type and a constant interest rate. More sophisticated price processes or interest rate assumptions would have more complicated test output, but the overall qualitative nature of the test output would not change much. If the price of the asset is denoted by X and the constant risk-free rate is denoted by r , then the martingale test at horizon T is based on the following expression:

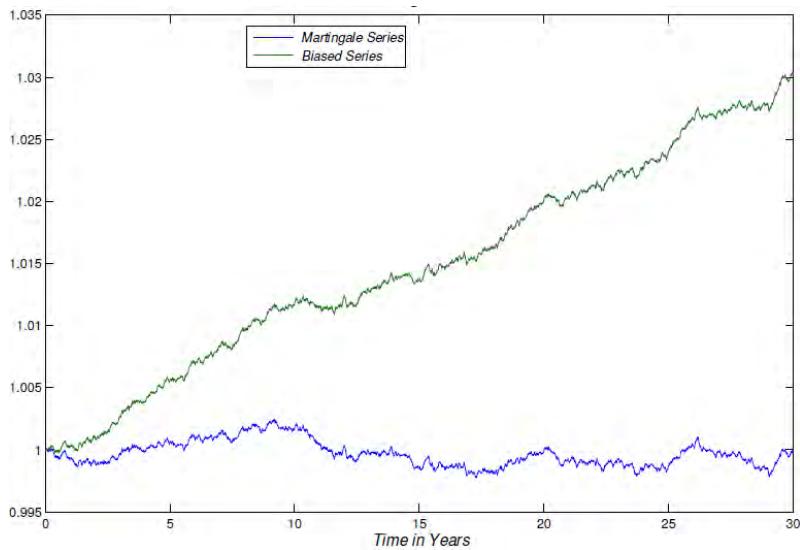
$$X_0 = e^{-rT} \frac{1}{M} \sum_{i=1}^M X_T(i).$$

We assume that $X_0 = 1$ so that we are checking that the test statistic is equal to 1. It is customary to generate the test statistic for all time periods to the largest horizon tested, and for our example, we run the test over a 30-year horizon. Consequently, the charts we are viewing show the path of martingale test statistics for all periods up to 30 years. If the martingale test is passed, each of these charts will tend to be located around 1 for all time points. Systematic

deviations from 1 indicate that the discounted asset price process is not a martingale. However, as we will observe, volatility makes it harder to diagnose test results. In each of our examples, the size of the risk-neutral scenario set is 10,000. For each example, we have generated a true risk-neutral series and an accompanying biased series. The biased series is based on exactly the same random draws as the true risk-neutral series, so the charts show the way in which the bias manifests itself relative to a series that would pass the martingale test.

The results in Figure 10.1 are typical for a moderate-volatility asset.

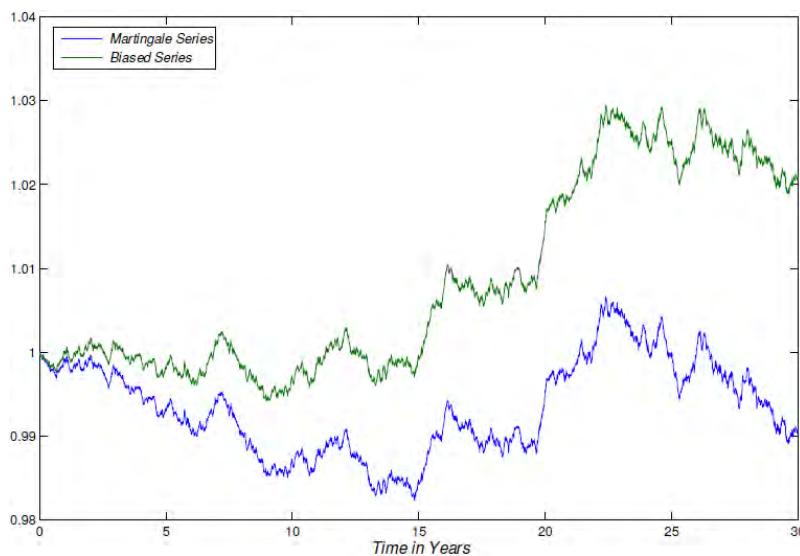
Figure 10.1
TYPICAL TEST RESULTS FOR MODERATE-VOLATILITY ASSETS, $\sigma = 0.08$



Source: Conning Inc.

The results in Figure 10.2 are typical for a high-volatility asset. The true martingale series may deviate from 1, but it should return to near this level over a sufficiently long horizon.

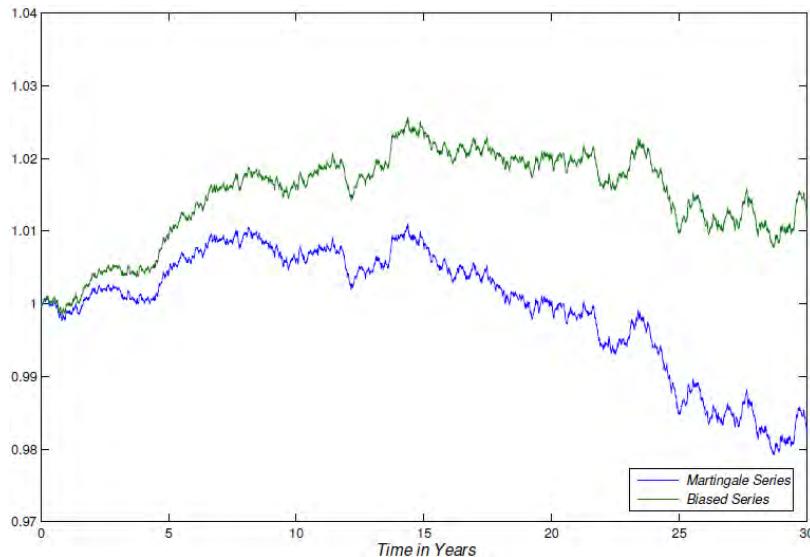
Figure 10.2
TYPICAL TEST RESULTS FOR HIGH-VOLATILITY ASSETS, $\sigma = 0.20$



Source: Conning Inc.

The results in Figure 10.3 illustrate an example that can create doubts when testing a high-volatility asset. The true martingale series appears to behave less like a martingale than the biased series. The reason for this is that this scenario set happens to have relatively unusual draws and the biased series tends to compensate for this. This shows why one should run the martingale test a few different times, even with a fairly large number of scenarios.

Figure 10.3
INCONCLUSIVE TEST RESULTS FOR HIGH-VOLATILITY ASSETS



Source: Conning Inc.

10.3 WHAT IS THE PURPOSE OF RISK-NEUTRAL SCENARIOS?

As we have already noted, risk-neutral scenarios arise in the connection between future cash flows and the current price of an asset. Roughly speaking, the connection is that the current price of an asset is the expected discounted value of the future cash flows associated with that asset. Ideally, our ESG models permit us to evaluate this expectation in closed form for a wide range of assets. However, complicated valuation problems will lack closed formulas.

Indeed, there are many important insurance applications for which the risk-neutral expectation that gives the price of an insurance liability cannot be readily evaluated. For example, an insurance company may offer an investment guarantee that depends on the performance of an investment based on a portfolio of several equity indices. The investment guarantee might depend on the history of the index returns, so that a sampling of the evolving equity indices figures in the resulting guarantee liability.

For illustration, let us suppose that this results in an expression of the following form:

$$V_T = \varphi(S_{t_1}, S_{t_2}, \dots, S_{t_N})$$

where S_{t_j} denotes the values of the equity indices at time t_j and $t_N = T$. The price of this uncertain cash flow is equal to

$$E \left[\exp \left(- \int_0^T r_u du \right) \varphi(S_{t_1}, S_{t_2}, \dots, S_{t_N}) \right].$$

This expression is difficult to evaluate in general, but it does provide a precise formula that represents the price of the liability. The accurate evaluation of this equation can be an intricate matter. In theory, if we are able to generate a risk-neutral scenario set for the model, then we can compute the expression under the expectation in this equation for each scenario and average the results to get the expected value that gives the price. This procedure is commonly

referred to as Monte Carlo simulation, and it is through situations like the one we have just described that Monte Carlo simulation naturally enters into the discussion.

We emphasize that this equation is a theoretical consequence of arbitrage-free pricing and is a statement about the relationship between risk-neutral scenarios and the price of a cash flow stream. The decision to compute this equation by Monte Carlo simulation depends on whether or not one is able to efficiently generate a large risk-neutral scenario set and whether alternative methods might be faster or more accurate. It is the application of risk-neutral scenarios to general Monte Carlo simulation that is the primary reason for considering risk-neutral scenarios. Alternative numerical methods such as Fourier inversion or finite difference methods may also be applied to compute this equation. Regardless of the numerical method used, the evaluation of this equation should not be confused with the basic notions of risk-neutral scenarios.

The idea of risk neutral is useful because it is intimately related to the pricing of assets and more general cash flow streams, and it provides a practical framework for computing prices. It is also useful to be able to parameterize a model to exhibit qualitative behavior that is consistent with market norms. This latter criterion is about getting output from the model to line up with real-world values. A natural question to ask is whether one can assess the reasonableness of risk-neutral output by comparing it with historical data. The answer is yes and no. One may say yes in that reasonable risk-neutral output will give prices for securities that are close to observed market prices. One must say no in that the averages of a risk-neutral variable cannot be benchmarked against historical data. Indeed, risk-neutral risk variables often have means that would be unreasonable in a real-world simulation context. If one obtains risk-neutral risk interest rate scenarios where some interest rates are 100%, is that OK? Probably not; it most likely reflects a model misspecification. There is a link between real-world and risk-neutral interest rates through the market prices of interest rate risk (and thus term premiums), as we observe in Chapter 11, and this link should provide a break on extreme divergence between real-world and risk-neutral scenarios.

When an ESG is utilized for risk-neutral simulation, the focus of the calibration will be on matching equity and swaption volatility surfaces and the reproduction of derivative market prices in general. The idea is that the ESG should be calibrated to market price data as tightly as possible, so that the pricing of insurance liabilities and other contingent claims that are not priced in the market will have prices that reflect liquid market prices as accurately as possible. Ensuring the calibration matches the initial yield curve also is an important consideration.

Risk-neutral scenarios are used to value cash flows. In cases where the cash flows that we seek to value are deterministic, risk-neutral scenarios are not needed, since the value of these cash flows can be computed by simple discounting. If we were to perform a risk-neutral Monte Carlo simulation to value these deterministic cash flows, the value we get back would agree with that obtained by simple discounting. The instances for which risk-neutral scenarios are needed are those for which one has complicated cash flows that depend upon stochastic financial variables. We then apply a risk-neutral scenario set to capture the covariation in the cash flow dependency with the key financial variables, in order to price the cash flow. Interest-sensitive insurance liabilities are a classic example of this paradigm. Market-consistent embedded value (MCEV), variable annuity risk management (i.e., exposure assessment and hedging) and the pricing of exotic cash flows are all examples in which risk-neutral scenarios are applied.

10.4 DECIDING BETWEEN REAL-WORLD AND RISK-NEUTRAL SCENARIOS

The general distinction is that risk-neutral scenarios are used for pricing and real-world scenarios are used to assess risk. Banks that trade derivatives are primarily interested in the price of the derivatives they are buying and selling, and risk-neutral scenarios are often their primary focus. Risk-neutral scenarios are typically used when one must price cash flows for which closed formulas are not available, and the risk-neutral scenarios are applied to numerically compute prices by Monte Carlo simulation.

Real-world scenarios are applicable for measuring returns, volatilities and other risk exposures. Risk-neutral scenarios are applicable for pricing cash flows. Real-world scenarios are calibrated to historical benchmarks or clients' own views. Risk-neutral scenarios are calibrated to pricing data, such as swaptions and index options.

The difference between risk-neutral scenarios and real-world scenarios is not the individual scenarios themselves; it is the probability of those scenarios occurring. Theoretically, all of the scenarios in a risk-neutral scenario set are real

world and vice versa. A risk-neutral scenario set weights the bad scenarios more than a real-world scenario set does, precisely because investors fear these scenarios. Risk-neutral scenarios weight adverse scenarios more heavily than real-world scenarios when pricing a security. Risk-neutral scenarios often look unreasonable if viewed relative to historical interest rates or returns. This is largely because risk-neutral scenarios embed the risk premiums necessary to get appropriate market prices.

10.5 WHEN ARE REAL-WORLD AND RISK-NEUTRAL SCENARIOS USED TOGETHER?

Real-world and risk-neutral scenarios will often be applied together in applications that combine risk management metrics with pricing. How does this arise? We have noted that risk-neutral scenarios are used for pricing and real-world scenarios are used to measure risk. The process of measuring risk almost always involves looking at returns, and returns depend on changes in price. One might therefore be tempted to assert that real-world scenarios must certainly directly employ risk-neutral scenarios if they are to produce any meaningful measure of risk. Such an assertion is not true, because ESGs are typically chosen so that the prices of key assets like treasury bonds, equities and corporate bonds are available in closed form. When the prices of key assets are available in closed form, these prices can be simulated off of the real-world dynamics for the economy, and risk-neutral scenarios are not required.

Risk-neutral scenarios will enter the picture when the real-world risk measurement exercise requires prices for liabilities or assets that do not have closed-form formulas. One purely financial example that occurs in practice is managing portfolios of mortgage-backed securities (MBSs). MBSs have complicated path-dependent cash flows and thus do not have convenient closed-form formulas for their prices.¹ The risk management and hedging of variable annuities is a standard example in which real-world and risk-neutral scenarios are applied together. Policyholder behavior and the availability of many investment products make it difficult to price the insurance liabilities associated with the product. An application such as this typically involves nested stochastic (i.e., inner loops) where the state of the world or node is simulated under the real-world measure, and cash flows are priced in that state of the world using risk-neutral scenarios projected forward from the current node of the simulation. In this case, the real-world simulation is used to assess the overall risk of the variable annuity book and to measure the effectiveness of the hedging strategy. Since the liabilities associated with the variable annuity book are complicated, closed-form formulas are not available, and risk-neutral scenarios must be used to price the variable annuity book at each node of the real-world simulation.

Inner loops are computationally intensive. If the pricing that the inner loop is being used for depends on a small number of financial variables, then it may be possible to precompute the prices and simply feed the relevant real-world output the precomputed prices. Unfortunately, many practical insurance problems cannot be simplified in this way. Therefore, other ways to avoid inner looping have been developed. Two important examples are replicating portfolios and least squares Monte Carlo, the details of which are beyond the scope of this publication.

10.6 CHAPTER SUMMARY

1. Risk neutral has a very specific meaning, which is inherently technical in nature but has an important practical manifestation that is the basis for generating pricing scenarios.
2. Risk-neutral scenarios are typically used when one must price cash flows for which closed formulas are not available and the risk-neutral scenarios are applied to compute prices. This situation is usually thought of in the context of Monte Carlo simulation. Although Monte Carlo simulation may be used to generate the risk-neutral scenarios, there is no other link between Monte Carlo and risk neutral. Market-consistent embedded value (MCEV), variable annuity risk management (i.e. exposure assessment and hedging) and the pricing of complicated cash flows are all examples in which risk-neutral scenarios are applied.
3. The martingale test is a check to confirm that a set of scenarios are risk neutral. The martingale test is often required by regulators. The implementation of the martingale test requires a careful assembly of the relevant risk-neutral variables. Volatile asset returns such as equity returns can produce martingale test results that are not perfect passes, even though the scenarios may be legitimate risk-neutral scenarios.

¹ Certain simplifying assumptions can be made to promote faster computation times, but sophisticated prepayment functions, such as those used by trading institutions, will usually require a risk-neutral approach.

4. Real-world scenarios are applicable for measuring returns, volatilities and other risk exposures. Risk-neutral scenarios are applicable for pricing cash flows.
5. Real-world scenarios are parameterized to reflect historical benchmarks or clients' own views. Risk-neutral scenarios are calibrated to pricing data such as swaptions and index options.
6. Risk-neutral scenarios often look unreasonable if viewed relative to historical interest rates or returns. This is largely because risk-neutral scenarios embed the risk premiums necessary to get appropriate market prices.
7. Risk-neutral scenarios have a reputation of being difficult. The addled view that is common among newcomers is largely the result of failing to understand the fundamental differences between real-world and risk-neutral scenarios.
8. Real-world and risk-neutral scenarios will often be applied together in applications that combine risk management metrics with pricing. Such applications typically involve stochastic on stochastic (i.e., inner loops) where the state of the world or node is simulated under the real-world measure and cash flows are priced in that state of the world using risk-neutral scenarios projected forward from the current node of the simulation. The risk management and hedging of variable annuities is a standard example in which real-world and risk-neutral scenarios are applied together. In this case, the real-world simulation is used to assess the overall risk of the variable annuity book and to measure the effectiveness of the hedging strategy. Since the liabilities associated with the variable annuity book are complicated, closed-form formulas are not available, and risk-neutral scenarios must be used to price the variable annuity book at each node of the real-world simulation.

10.7 CHAPTER REFERENCES

The material in Chapter 10 is an application of the ideas of Chapter 9. The references from Chapter 9 may be consulted for the material of this chapter. Additional sources include the following.

- Bingham, Nicholas H., and Rüdiger Kiesel. 2004. *Risk-Neutral Valuation: Pricing and Hedging of Financial Derivatives*. 2nd ed. New York: Springer.
- Hunt, P. J., and J. E. Kennedy. 2004. *Financial Derivatives in Theory and Practice*. Revised ed. New York: John Wiley & Sons.

Chapter 11: Default-Free Interest Rate Models

As a key component of an ESG, the default-free interest rate model has the primary purpose of generating the prices of zero-coupon bonds for the economy. Having knowledge of the prices of zero-coupon bonds allows the construction of yield curves and the pricing of all default-free interest-rate contingent cash flows. Practitioners often refer to the model as generating the *term structure of interest rates*, or *term structure* for short.

Section 11.1 discusses important properties of robust interest rate models. For example, they should model the economy so no arbitrage-trading opportunities are available, which permits utilization of the general mathematical framework of arbitrage-free models, discussed in Chapter 9. An ESG should begin the simulation with the initial model yield curve in agreement with the market observed yield curve. Another consideration addresses the phenomenon of negative interest rates, which has of late become a significant challenge to interest rate models.

Section 11.2 discusses binomial interest rate models, which are based on a discrete time series and have been applied since the earliest days of interest rate modeling. Binomial models offer a computationally flexible approach for pricing a wide variety of cash flow structures. Multiperiod dynamics for the binomial model are described as a multiplicative binomial model.

Section 11.3 begins a discussion of a continuous-time single-factor model. Going from discrete periods, such as considered in the binomial model, to continuous time requires a more difficult level of mathematics, but this provides a better model of interest rate dynamics with greater computational flexibility. The basic building block for these factor models is a stochastic differential equation for the short rate. The other input to the model is the market price of risk process. Examples of these models are the Gaussian or Vasicek model and the Cox-Ingersoll-Ross (CIR) model.

Section 11.4 provides some information on more advanced models. The section includes a more generalized discussion of affine models, beginning with a more detailed discussion of a two-factor continuous-time (affine) model. The section then discusses models that extend beyond affine continuous-time models to classes described as quadratic interest rate models, shadow interest rate models, threshold interest rate models, regime-switching models and macro finance models (specifically the dynamic Nelson-Siegel model). An advantage of the DNS framework is that a wide range of factor dynamics can be introduced. Furthermore, the model is often integrated into a macro-finance framework, where the factor dynamics are directly linked to key macro variable dynamics. A drawback of the model is that it is not suitable for derivatives pricing and is therefore a tool for real-world applications.

11.1 IMPORTANT PROPERTIES OF ROBUST INTEREST RATE MODELS

Realistic models of interest-rates should have many important properties. For example, short rates tend to be more volatile than long rates. An ESG should also model the economy so no arbitrage-trading opportunities are available, and this perspective permits us to bring the general mathematical framework from Chapter 9 (arbitrage-free models) to bear on the problem. This framework may then be used to accommodate some of the stylized facts (as discussed in Chapter 6) that we believe a reasonable term structure model ought to possess.

A robust interest rate model should be capable of producing realistic interest rate scenarios and distributions in real-world applications and efficient pricing of cash flows in risk-neutral settings. The model should have a tractable calibration algorithm that can be used with market data to accurately reflect the relevant features of the risk management application. The accuracy of the model should be examined using a careful validation procedure. One might think that the key to building a “good” model of the term structure is to model the correct observable economic variables. Should not variables such as short rate, long rate and the inflation rate figure in the dynamics of interest rates somehow?

While this is true, it is disappointing to learn that models that attempt to explicitly model these variables tend to perform poorly; the models are normally misspecified because the ways in which these variables interact to produce interest rates are too complex to be captured in a “simple” model. In practice, *latent variable models* (models in which unobservable factors are used to explain fluctuations in interest rates) tend to perform better in practice than

do such explicit economic models. Often one can get better performance from a model that uses latent (i.e., unexplained) factors than from trying to model observable factors directly.

The basis for the construction of arbitrage-free interest rate models is the observation that zero-coupon bonds are traded securities. As traded securities, zero-coupon bonds are subject to the technical restrictions imposed by arbitrage-free pricing, discussed in Chapter 9. These technical restrictions also serve as a framework for developing interest rate models where the starting point is the requirement that the dynamics of the discounted zero-coupon bonds behave as martingales under a risk-neutral measure. Therefore, the martingale condition will be a part of any arbitrage-free interest rate model.

In most applications of an ESG, it is important that the ESG begins the simulation with the initial model yield curve in agreement with the market observed yield curve.¹ Affine models will typically agree with the market observed yield curve at a number of points equal to the number of factors employed in the model. One method to adjust the yield curve at positions other than these points is to use a deterministic shift, as explained in detail by Dybvig(1997). In extreme environments there is a practical issue in that the shift will propagate forward through time and can cause unwanted distortions to the simulation at a later time at a shorter maturity on the yield curve. An alternative to a deterministic shift is to allow time dependent parameters in the factor processes, as was done in Hull and White (1990). The differences between these two methods are material. The use of time-dependent coefficients can also be used to induce a temporary change in the dynamics of the underlying state variables. However, the use of time-dependent coefficients effectively introduces many extra parameters to the model.

From time to time in the literature, there seems to be considerable confusion about what it means for an interest rate model to not match the initial yield curve. Some authors have asserted that not exactly producing the initial yield curve violates the condition that the model is arbitrage-free. This is not theoretically true, but it does present some practical problems in setting up an initial bond portfolio with a known market value. Models that are constructed in the presence of an equivalent martingale measure are arbitrage-free, but they may not be capable of being calibrated exactly to a given yield curve. This is a misspecification error rather than an arbitrage. So it is possible for an arbitrage-free model to have difficulty capturing some initial yield curves.

11.1.1 NEGATIVE INTEREST RATE

The recent emergence of significant negative interest rates raises challenges and questions. Some interest rate models do not permit negative interest rates. This typically arises because of the implications of the types of stochastic processes used to construct the interest rate model, two common examples being that interest rates are log normally distributed or non-central chi-square distributed. Negative interest rates can still be allowed for in such models if a shift is carefully applied to the model dynamics. The constant shift moves all interest rates relative to their positions in the model prior to applying the constant shift. A constant shift also preserves the shape of the distribution of interest rates. The method has the advantage that it is easy to apply. The method has the disadvantage that the shifted distribution of interest rates may not have a left tail that is consistent with history or the calibration view.

The deeper modeling questions, and ones that are under very active research, focus on the issue of what classes of interest rate models are suitable for capturing the interest rate experience of the past decade. Models that are specifically geared to the low interest rates seen globally post financial crisis often refer to the zero-lower bound (ZLB). We now know that there is no hard zero-lower bound, but the important idea associated with ZLB is that interest rates are not likely to take on negative values of large absolute value.²

¹ Some models will automatically produce the market yield curve, two examples being LIBOR market models and Heath-Jarrow-Morton (HJM) models. LIBOR market models were developed for derivative pricing and may not exhibit realistic interest rate behavior over long simulation horizons. HJM models are not generally finite-dimensional if a wide range of volatility dynamics is required. Therefore, one may not necessarily select models that easily accommodate the initial yield curve, as other performance considerations play a role.

² An interest rate of -1.0% is already near the extreme of what is expected. An interest rate of -7.0% seems implausible and would not be wanted as ESG model output.

11.1.2 ARE NEGATIVE INTEREST RATES ARBITRAGE?

If an interest rate model admits negative interest rates, there is a temptation to immediately conclude that the model is not arbitrage-free. After all, why not just hold cash in a mattress or a bank vault instead of purchasing bonds for more than the sum of their future cash flows? However, the arbitrage-free framework requires that all tradable assets be specified at the outset, and all traded assets are then subject to the no-arbitrage conditions. The asset that is specified for holding cash is the money market account, and this “earns” the risk-free interest rate. Market participants have no other ways to hold cash. A “cash in the mattress” investment was not part of the specified assets.³ These days, one often hears talk from some economist about how the government should abolish cash so that it can enforce negative interest rates if the need should arise. The actual economy has market frictions that are not included in a theoretical arbitrage-free framework. There are various ways to square the implications of the theory with the range of issues that are emerging in the actual economy.

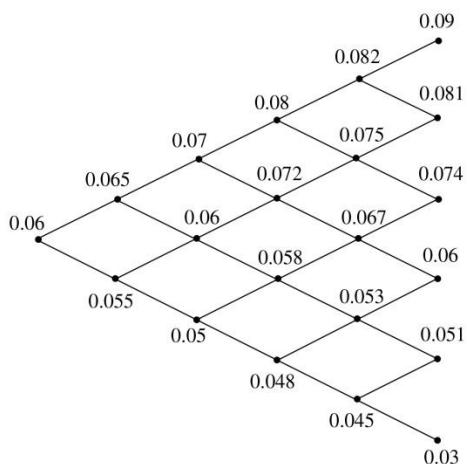
11.2 BINOMIAL INTEREST RATE MODELS

Binomial interest rate models have been applied since the earliest days of interest rate modeling. The first applications were to valuing contingent cash flows. Binomial models offer a computationally flexible approach for pricing a wide variety of cash flow structures. Binomial models may be understood from elementary first principles, in contrast to the more technical issues that must be dealt with in continuous-time models. Binomial models can be linked to continuous-time models through limiting processes and their connection to finite difference solutions of partial differential equations.

A typical example of a binomial interest rate model is shown in Figure 11.1. The model consists of a lattice (also called tree) structure that is used to specify the development of the one-period interest rates, the values of one-period interest rates at each node of the lattice, and an assumption about the probability of moving up or down from each node. Figure 11.1 is a five-period lattice model.

Figure 11.1

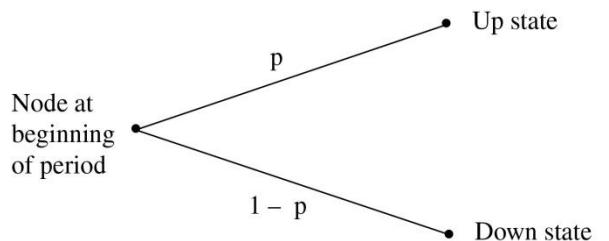
AN ILLUSTRATIVE BINOMIAL INTEREST RATE MODEL



Let us develop the basic formulas for binomial interest rate models. We will take the following argument and use it to understand the model dynamics at every node of the interest rate tree. In the following, we are at some fixed but otherwise generic node in the model. The same set of arguments we are about to make will apply to every node in the lattice. A picture for our generic node is as shown in Figure 11.2.

³ If one tries to define a “cash in the mattress” asset, one will find that the no-arbitrage conditions force it back to a money-market-type instrument.

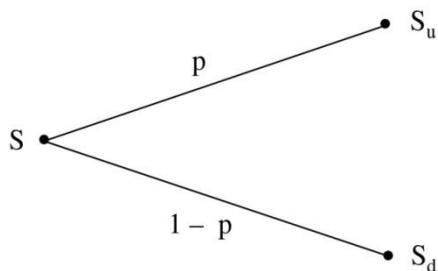
Figure 11.2
BINOMIAL BRANCHING AT GENERIC NODE



Model branches to one of two possible states at end of period

We use p to denote the conditional probability of the up state occurring at the end of the period. If S is the price of a security at the beginning of the period, the security will have a price S_u if the up state prevails at the end of the period, and a price S_d if the down state prevails at the end of the period, as represented by Figure 11.3. One may think of the security as a bond, but it is intended to represent a generic set of cash flows.

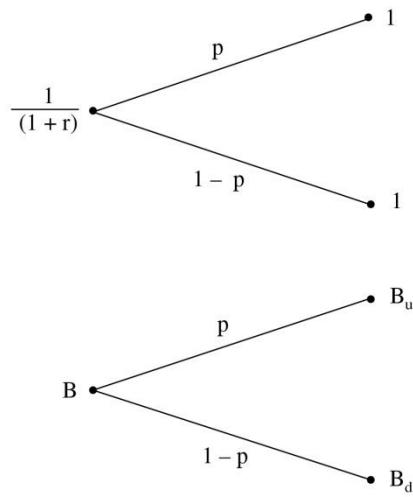
Figure 11.3
GENERIC PRICE DYNAMICS OF A SECURITY



Let the one-period rate (also called the short rate) in effect at the beginning of the period be denoted by r . The value of r depends on which node we are at in the lattice, but we will suppress this dependence in our notation. We now derive the most important implication of the absence of arbitrage.

We suppose that available for trade is the one-period bond and a risky bond, which we will denote by B . The requirement that the bond be risky means simply that $B_u \neq B_d$; i.e., the future value of the bond at the end of the period is unknown. The prices of these two assets will evolve as shown in Figure 11.4.

Figure 11.4
GENERIC PRICE DYNAMICS OF A SECURITY



Now consider a third generic security denoted by S , the price of which evolves, as we have already noted, as in Figure 11.3. We are interested in relating the price of this security to the prices of the traded and priced securities. To this end, consider a portfolio of the one-period bond and the risky bond, in which ϕ units of the one-period bond and θ units of the risky bond are held:

$$\text{Cost of portfolio} = \frac{\phi}{1+r} + \theta B$$

$$\text{Payoff in the up state} = \phi + \theta B_u$$

$$\text{Payoff in the down state} = \phi + \theta B_d$$

We seek values for ϕ and θ such that the payoffs from this portfolio in each of the up and down states is exactly equal to the payoff from the generic security S we are considering. Mathematically, we are seeking ϕ and θ such that

$$\begin{aligned}\phi + \theta B_u &= S_u \\ \phi + \theta B_d &= S_d\end{aligned}$$

By simple linear algebra, we find that there is such a portfolio with

$$\begin{aligned}\theta &= \frac{S_d - S_u}{B_d - B_u} \\ \phi &= S_u - B_u \frac{S_d - S_u}{B_d - B_u}\end{aligned}$$

The only requirement to make the algebra work is that $B_u \neq B_d$. The cost of this portfolio is equal to

$$\phi \frac{1}{1+r} + \theta B = \left(S_u - B_u \frac{S_d - S_u}{B_d - B_u} \right) \frac{1}{1+r} + \frac{S_d - S_u}{B_d - B_u} B$$

This expression may be simplified to

$$\text{Cost of portfolio} = \frac{1}{1+r} \left[S_u \frac{B_d - B(1+r)}{B_d - B_u} + S_d \frac{B(1+r) - B_u}{B_d - B_u} \right]$$

Thus, we have determined a portfolio of the one-period and risky bonds that has exactly the same payoffs as our generic security S . Therefore, **if there is no arbitrage, this portfolio must cost the same as the generic security.** Indeed, if they sold for different prices, then we could buy one and short the other, making an instant profit. This piece of reasoning is the most common application of the principle of no arbitrage.

The price of the generic security today was denoted S . Since this is equal to the cost of the portfolio, we have shown that

$$S = \frac{1}{1+r} \left[S_u \frac{B_d - B(1+r)}{B_d - B_u} + S_d \frac{B(1+r) - B_u}{B_d - B_u} \right]$$

It is straightforward to check that if there is no arbitrage; then $B(1+r)$ must be between B_u and B_d . The argument is that if this were not the case, then the one-period bond must dominate the risky bond or vice versa—either of which would result in an arbitrage.

Let us define

$$\pi = \frac{B_d - B(1+r)}{B_d - B_u}$$

Then π is a number such that $0 < \pi < 1$ and is called a risk-neutral (or martingale) probability.

Since S was assumed to be an arbitrary security, we have shown that all securities satisfy the equation

$$(11.1) \quad S = \frac{1}{1+r} [\pi S_u + (1-\pi) S_d]$$

This is a very important result. It says that the price of a security is equal to the discounted expectation of its payoffs, where we discount at the short rate and the expectation is taken with respect to the risk-neutral probabilities. This is one of the most important ideas in financial economics. Note also that the actual probability p **does not figure in the pricing of the generic security.**⁴

In the context of binomial interest rate models, Equation 11.1 means that if we specify the one-period interest rate and binomial risk-neutral probability at every node of the lattice, then we have a model that can be used to price contingent cash flows. In many models, the binomial probability is assumed to be 1/2. The manner in which the one-period interest rates are specified is the key distinguishing feature of different binomial interest rate models.

We give an illustrative example of the specification of a binomial interest rate model and a simple example to illustrate the basic application of the model.

We consider a multiplicative binomial model as defined by the local dynamics for each node shown in Figure 11.5.

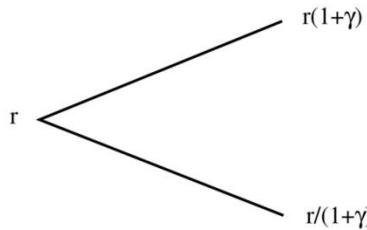
⁴ Providing that S is a risky asset (i.e., $S_u \neq S_d$, which is required for the algebra to be defined), we can use this formula to solve for π in terms of the asset.

Performing the simple algebra one finds

$$\pi = \frac{S_d - S(1+r)}{S_d - S_u}$$

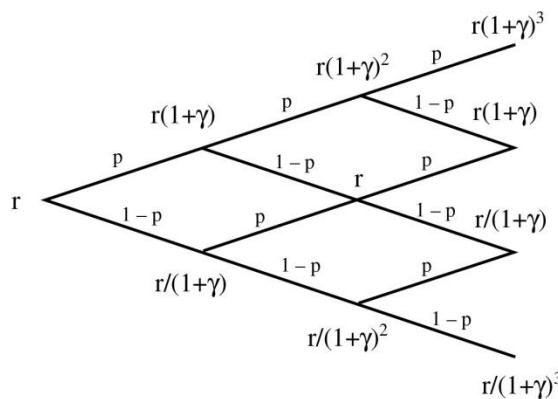
and the right side of the expression will be the same for all risky assets. Thus, knowing the payoffs of a risky asset and its price today permits us to solve for π .

Figure 11.5
LOCAL DYNAMICS AT EACH NODE OF THE BINOMIAL LATTICE



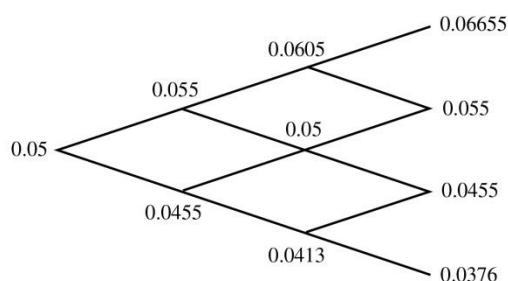
At each node in the lattice, the one-period interest rate can branch up or down to a value that is determined from the value of the one-period rate at the start of the period, and a parameter γ . When the local dynamics are put together, we obtain the multiperiod interest rate tree shown in Figure 11.6.

Figure 11.6
MULTIPEIOD DYNAMICS FOR THE BINOMIAL MODEL



The tree, together with the parameter value p , is called the multiplicative binomial model. In Figure 11.7, we construct an illustrative three-period multiplicative binomial interest rate tree for the parameter values $p = 0.5$; $r = 0.05$; $\gamma = 0.1$.

Figure 11.7
ILLUSTRATIVE MULTIPEIOD BINOMIAL MODEL



To price cash flows in this model, you simply discount them back through the lattice in accordance with the relationship noted in Equation 11.1. For example, suppose we want to price a three-period zero-coupon bond maturing for a unit amount. A three-period zero-coupon bond has cash flows of 1 at time three, and no cash flows elsewhere, as shown in Figure 11.8.

Figure 11.8
CASH FLOWS FROM A THREE-PERIOD ZERO-COUPON BOND MATURING FOR A UNIT AMOUNT

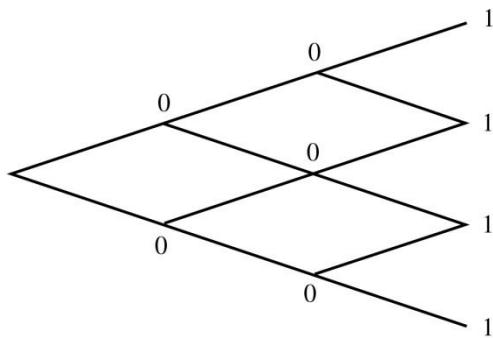
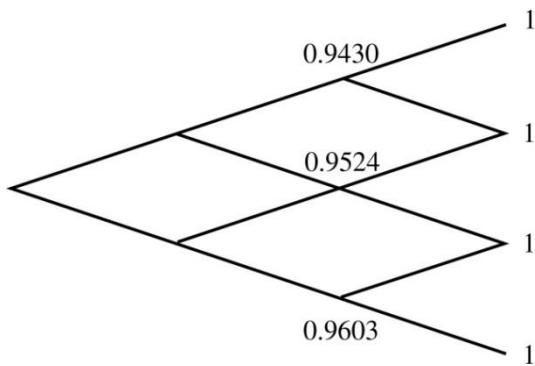


Figure 11.9
RECURSIVELY PRICE AT EACH NODE



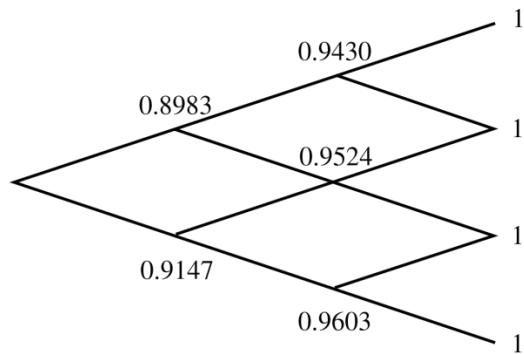
We value the cash flows recursively as follows (see Figures 11.9 through 11.11). First find the prices at period two. The following example shows the top node at period two:

$$0.9430 = \frac{1}{1.0605} [0.5(1) + 0.5(1)]$$

The other values of 0.9524 and 0.9603 are computed similarly as the discounted expected payments of the cash flows at time 3 that emanate from the node we are computing. When these recursion are completed back to time 0 we find that the price of the three period zero-coupon bond maturing for a unit amount is 0.8633. More complicated contingent cash flows can be priced, but the procedures remain similar.

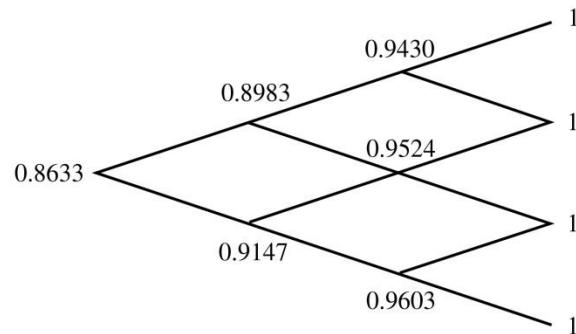
The model may be generalized to different branching structures, such as a trinomial mechanism, and to a multifactor context.

Figure 11.10
RECURSIVELY PRICE AT EACH NODE



$$0.8983 = \frac{1}{1.055} [0.5(0.9430) + 0.5(0.9524)]$$

Figure 11.11
RECURSIVELY PRICE AT EACH NODE



$$0.8633 = \frac{1}{1.05} [0.5(0.8983) + 0.5(0.9147)]$$

There is a large literature on binomial interest rate models. See van der Hoek and Elliott (2006), Tuckman (2002), Lyuu (2004), Ho and Lee (2004), and Black and Karasinski (1991) for additional information.

11.3 CONTINUOUS-TIME SINGLE-FACTOR MODELS OF THE TERM STRUCTURE

Continuous-time interest rate models figure prominently in ESGs because they have been extensively researched, and they offer a flexible and tractable framework for the simulation of realistic interest rate dynamics and the pricing of contingent cash flows. Continuous-time interest rate models were the first class of interest rate models that were extensively researched in the late 1970s, with the development of the Cox-Ingersoll-Ross and Vasicek models.

Continuous-time interest rate models can be based on a single factor or multiple factors. Single-factor models have the advantage of simplicity. They are easier to calibrate and simulate than their multifactor cousins. However, single-factor models have some important deficiencies that can render them inappropriate for some applications. For example, since a single-factor model has one source of uncertainty, changes in interest rates of different tenors are

perfectly correlated. Also, the yield curves associated with a single-factor model are completely determined by one tenor, so there is only one possible yield curve associated with a given level for a specific tenor interest rate. In other words, for a single-factor interest rate model, there is one and only one yield curve the model can generate that has a three-month interest rate of 1%. A careful examination of historical yield curves shows that there have been many different shapes of yield curve associated with a three-month interest rate of 1%, so it is fair to say that a single-factor interest rate model cannot produce a wider range of yield curve shapes. Therefore, some applications will be fine with the use of a single-factor interest rate model, while others will need the more sophisticated multifactor framework.

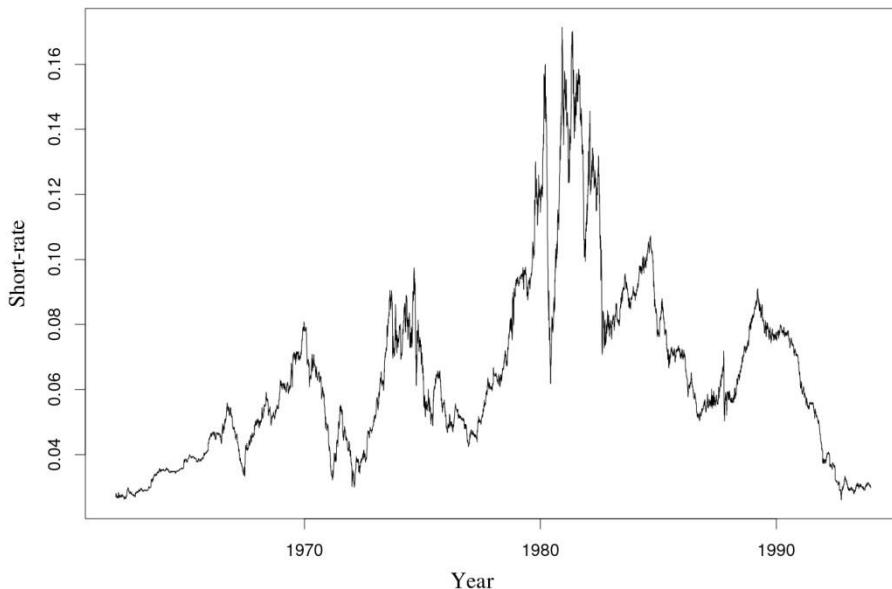
Most of the key ideas in continuous-time interest rate modeling can be illustrated in the single-factor context. For this reason, we restrict most of our discussion to this class of models.

11.3.1 Fine Details of One-Factor Continuous-Time Models

In going from discrete to continuous time, one must accept a more difficult level of mathematics. In return for the increased complexity, one is better able to model interest rate dynamics and acquire a model with somewhat greater computational flexibility.

The basic building block for the one-factor continuous-time model is a stochastic differential equation for the short rate. It is customary to denote the short-rate process by r_t . Historical proxy data for the short rate is exhibited in Figure 11.12. This figure is a graph of the three-month U.S. Treasury rate based on the Federal Reserve Bank's H15 data set.

Figure 11.12
THE U.S. SHORT RATE



Source: St. Louis Federal Reserve, Federal Reserve Economic Data

The stochastic differential equation for the short-rate is assumed to be of the form

$$(11.2) \quad dr_t = \mu(r_t)dt + \sigma(r_t)dW_t$$

where W is a Brownian motion process.

For our purposes, the most significant properties of a Brownian motion process are as follows:

1. The process is continuous.
2. The increments of the process are independent.
3. $W_t - W_s$ is normally distributed with mean 0 and variance $t - s$.

We can gain some insight into the dynamics of Equation 11.2 by writing its “discretized” version as follows:

$$(11.3) \quad \begin{aligned} r_{t+\Delta t} - r_t &\approx \mu(r_t)\Delta t + \sigma(r_t)(W_{t+\Delta t} - W_t) \\ &= \mu(r_t)\Delta t + \sigma(r_t)\sqrt{\Delta t}\varepsilon_{t+\Delta t} \end{aligned}$$

where $\varepsilon_{t+\Delta t}$ denotes a standard normal random variable (i.e., normally distributed random variable with mean 0 and variance 1).

In practice, one seeks to find models for the short rate that reflect the empirical properties of interest rate time series. As one would expect, this is a challenging issue. Indeed, not only must we be able to formulate a mathematically tractable model, but we must be able to estimate the model parameters.

The other input to the model is the market price of risk process. This process governs the risk-return relationship for all bonds through the following general relationship:

$$\text{Market price of risk} = \frac{\text{Instantaneous expected return on bond} - r_t}{\text{Instantaneous standard deviation of bond return}}$$

The market price of risk process permits us to distinguish two versions of the interest rate model: “real world” and “risk neutral.” The real-world version is the one we use to simulate the empirical behavior of interest rates. The risk-neutral version is the one we use for pricing general interest rate contingent cash flows, derivatives, investment guarantees and so on.

A specification of the market price of risk (usually denoted λ_t) is equivalent to a specification of an equivalent martingale measure (usually denoted Q). However, we will avoid the more technical mathematical details and simply illustrate what effect the market price of risk has and how one computes it for the Vasicek model.

11.3.2 The Vasicek or Gaussian Model

One of the best-known models is the Gaussian model. In this model, the short-rate process is assumed to follow the following process:

$$dr_t = \alpha(\gamma - r_t)dt + \eta dW_t$$

where α , γ and η are constant. This model has mean reversion and constant short-rate volatility. The market price of risk process for this model is normally assumed to be a constant, which we will denote by λ .

The price of a discount bond at time t that matures for a unit at time T when the short rate is equal to r_t at time t is denoted by $P(r_t, t, T)$.

With these assumptions, one can compute a closed formula for the price of a discount bond as follows:

$$(VAS-PRICE) \quad P(r, t, T) = \exp(-rF(T-t) - G(T-t))$$

where

$$\begin{aligned} F(s) &= \frac{1 - e^{-\alpha s}}{\alpha} \\ G(s) &= sD - F(s)D + \frac{\eta^2}{4\alpha}F(s)^2 \\ D &= \gamma - \frac{\lambda\eta}{\alpha} - \frac{\eta^2}{2\alpha^2} \end{aligned}$$

Using the formulas provided by (VAS-PRICE), we can simulate most of the key aspects of the term structure needed for actuarial applications.

A more general version of this model, usually referred to as the Hull and White model, allows the parameters of the short-rate process to be time-dependent:

$$(11.4) \quad dr_t = \alpha(t) (\gamma(t) - r_t) dt + \eta(t) dW_t$$

The additional flexibility permits calibration to the initial yield curve and to derivatives prices in some risk-neutral pricing applications.

11.3.3 The Cox-Ingersoll-Ross (CIR) Model

Another popular model of the term structure is the Cox-Ingersoll-Ross (CIR) model. In this model, the short-rate process is assumed to follow the following process:

$$(11.5) \quad dr_t = \kappa(\theta - r_t)dt + \sigma\sqrt{r_t}dW_t$$

where κ , θ and σ are positive constants.

The dynamics in Equation 11.5 are assumed to be real world. This model has mean reversion and short-rate volatility that is proportional to the level of the short rate. The latter property is somewhat consistent with the volatility of the short rate that seems to be present in Figure 11.12, in that the volatility of the short rate appears to increase when the short rate is larger.

A property of this model is that we always have $r > 0$ for appropriate parameter restrictions. This used to be a requirement, but in recent times negative interest rates have become the norm.

The drift term $\kappa(\theta - r_t)$ is called mean reverting because the further r is from θ , the greater the pull this term imparts to bring r back to θ . For instance, suppose that $\theta = 0.06$ and r is currently equal to 0.09. The drift term is then equal to $\kappa(0.06 - 0.09) = -0.03\kappa$, a negative value, which will tend to make r decrease toward the mean level of 0.06.

Obviously, the larger the value of κ , the greater the effect of this term. Consequently, κ is referred to as the speed of the mean reversion.

The market price of risk process for this model is normally assumed to be a constant multiple of the square root of the short rate. We denote the constant by λ .

In the CIR model, the short-rate process is time homogeneous, meaning the drift and instantaneous volatility of the short rate do not depend on time t . Consequently, all discount bond prices are of the form

$$P(r, t, T) = P(r, T - t).$$

In other words, discount bond prices are of the form $P(r, \tau)$, meaning that bond prices are functions of the short rate and the remaining time to maturity.

With these assumptions, one can compute a closed formula for the price of a discount bond as follows:

$$(CIR-PRICE) \quad P(r, t, T) = A(T - t) \exp(-rB(T - t))$$

where

$$A(s) = \left(\frac{2\gamma \exp((\kappa + \lambda + \gamma)s/2)}{(\kappa + \lambda + \gamma)(\exp(\gamma s) - 1) + 2\gamma} \right)^{\frac{2\kappa\theta}{\sigma^2}}$$

$$B(s) = \frac{2(\exp(\gamma s) - 1)}{(\kappa + \lambda + \gamma)(\exp(\gamma s) - 1) + 2\gamma}$$

$$\gamma = \sqrt{(\kappa + \lambda)^2 + 2\sigma^2}$$

Using the formulas provided by (CIR-PRICE), we can simulate most of the key aspects of the term structure needed for actuarial applications.

11.3.4 SOME REPRESENTATIVE PARAMETER VALUES

It is possible to use a variety of statistical techniques to determine the model parameters for both the Vasicek and CIR models. Such methods include the following examples:

- Maximum likelihood
- Generalized method of moments
- Bayesian estimation

Later we will show one method for estimating the parameters for the Vasicek model: the method of maximum likelihood.

For the time being, we will use the following parameter estimates that were based on data from 1988–1993 by Chen and Yang (see Chen 1996). Any estimates of model parameters tend to exhibit substantial sensitivity to the period from which the data is drawn. In practice, one must be careful to estimate models across time periods for which the “economic regime” driving interest rates appear to be stable. The late 1970s and early ‘80s are an example of a change in regime—the Fed experiment based on money supply. Although regimes are an important part of economic data, single-regime models are not equipped to handle such drastic changes. More advanced models with “regime switching” are needed to estimate across periods with obvious shifts in regime.

The initial values of r_0 are typically taken from the Treasury T-bill rate at the start of the simulation period. The illustrative parameters from Chen (1996) are as shown below.

Parameters for the Vasicek Model

$$r_0 = 0.06$$

$$\alpha = 0.2456$$

$$\gamma = 0.0648$$

$$\eta = 0.0289$$

$$\lambda = -0.2718$$

Parameters for the CIR Model

$$r_0 = 0.06$$

$$\kappa = 0.2456$$

$$\theta = 0.0648$$

$$\sigma = 0.1499$$

$$\lambda = -0.1290$$

We may now use these parameter values to compare the behavior of yield curves for these two models and to generate some interest rate scenarios for each model.

Figures 11.13 and 11.14 show some representative yield curves from the Vasicek and CIR models for the preceding parameter values. In the following two figures, notice that the yield curves are indexed by the level of the short rate. This is characteristic of one-factor models.

Figure 11.13

SAMPLE YIELD CURVES FROM THE VASICEK MODEL AS r_0 VARIES

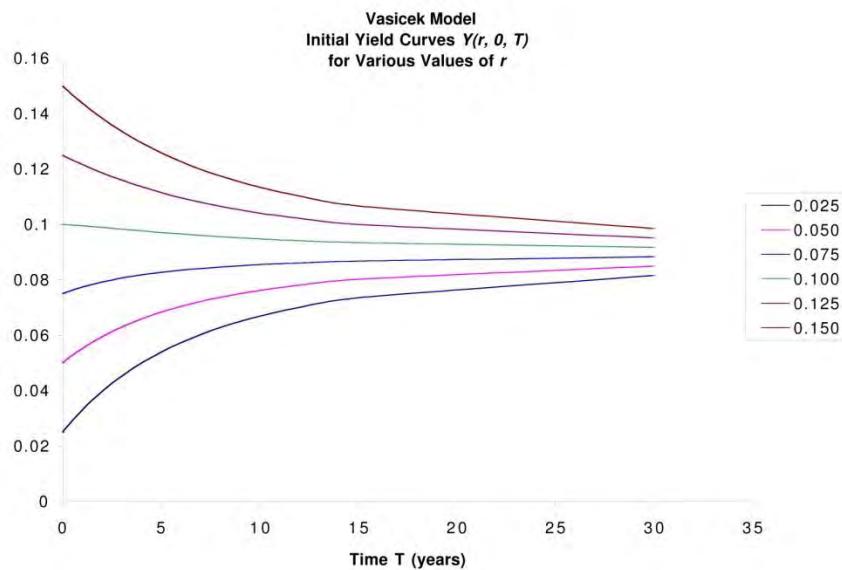
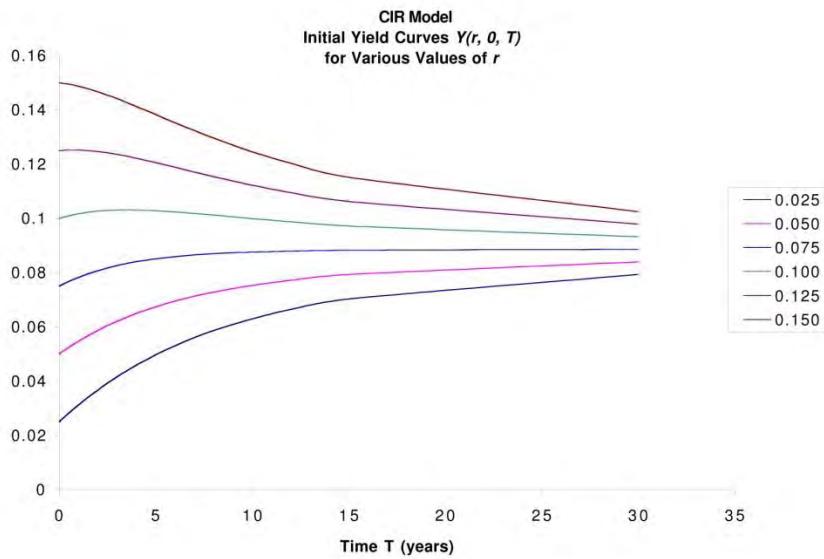


Figure 11.14

SAMPLE YIELD CURVES FROM THE CIR MODEL AS r_0 VARIES



11.3.5 MONTE CARLO VALUATION OF INTEREST RATE DERIVATIVES FOR THE VASICEK MODEL

The alternative to valuation by the use of partial differential equations is to use probabilistic techniques. The main advantage of the probabilistic viewpoint is the great flexibility and power of the Monte Carlo method. The primary

disadvantage of the probabilistic viewpoint is that one must understand the transformation of the short-rate process under the equivalent martingale measure.

We will give a concrete illustration of what is involved in the valuation of interest rate derivatives from the probabilistic viewpoint for the Vasicek model. As we know, the short-rate process for the Vasicek model is

$$dr_t = \alpha(\gamma - r_t)dt + \eta dW_t$$

Suppose that the market price of risk process is a constant denoted by λ . Then one can show that under the equivalent martingale measure, the short-rate has the dynamics

$$(TRANS) \quad d\tilde{r}_t = [\alpha(\gamma - \tilde{r}_t) - \eta\lambda]dt + \eta d\tilde{W}_t$$

where \tilde{W} is a Brownian motion. In other words, under the equivalent martingale measure, the dynamics of the short rate are very similar to the original short-rate dynamics except that we must adjust the drift of the short rate by a term involving the market price of risk. If we know how to make this adjustment for a particular model, then all risk-neutral expectations for that model can be computed or approximated.

Consider now an ordinary European call option with expiration date t , written on a discount bond maturing at time s ($s > t$) and with strike price K . The call option will pay the contingent payoff equal to

$$\max(0, P(r_t, t, s) - K)$$

at time t .

It can be shown that the price of the call option is given by the risk-neutral expectation

$$(EMM-FORMULA) \quad E^Q \left[\exp \left(- \int_0^t r_u du \right) \max(0, P(r_t, t, s) - K) \right]$$

Evaluating this expectation is equivalent to solving the partial differential equations we looked at earlier.

However, we can transform this expectation into something we can compute directly using the fact (TRANS). Indeed, substituting for r_t in (EMM-FORMULA) using (TRANS) yields the following formula for the price of the call option:

$$(ORDINARY) \quad E \left[\exp \left(- \int_0^t \tilde{r}_u du \right) \max(0, P(\tilde{r}_t, t, s) - K) \right]$$

It is possible to take advantage of properties of the normal distribution to compute a closed-formula solution for the price of the call option. However, one can also use approximate numerical methods (Monte Carlo) to obtain the call option price. This technique will work for all contingent claims and depends only on knowing the proper form of the transformed dynamics (TRANS).

We begin by dividing the interval $[0, t]$ into N pieces of equal length. Denote the implied discretization interval width by

$$h := \frac{t}{N}$$

Under the equivalent martingale measure, the short rate has the dynamics

$$d\tilde{r}_t = [\alpha(\gamma - \tilde{r}_t) - \eta\lambda]dt + \eta d\tilde{W}_t$$

which is discretized as

$$\tilde{r}_{t+\Delta t} - \tilde{r}_t = [\alpha(\gamma - \tilde{r}_t) - \eta\lambda]\Delta t + \eta \sqrt{\Delta t} \varepsilon_{t+\Delta t}$$

where $\varepsilon_{t+\Delta t}$ is normally distributed with mean 0 and variance 1. The discretized dynamics for the short rate can be rearranged as

$$\tilde{r}_{t+h} = \tilde{r}_t + [\alpha(\gamma - \tilde{r}_t) - \eta\lambda]\Delta t + \eta\sqrt{\Delta t}\varepsilon_{t+\Delta t}$$

from which a trajectory of short-rate values can be recursively generated from a starting value $\tilde{r}_0 \equiv r_0$ and a series of draws from the standard normal distribution.

We will use the equivalent notation $h = \Delta t$ for simplifying some expressions. This results in a trajectory of the short-rate process

$$\tilde{r}_h, \tilde{r}_{2h}, \tilde{r}_{3h}, \dots, \tilde{r}_{Nh}$$

Note that $Nh = t$.

We then approximate the random variable we want to compute the expectation of, namely

$$\exp\left(-\int_0^t \tilde{r}_u du\right) \max(0, P(\tilde{r}_t, t, s) - K)$$

by

$$\exp(-[\tilde{r}_h + \tilde{r}_{2h} + \dots + \tilde{r}_{Nh}]h) \max(0, P(\tilde{r}_{Nh}, t, s) - K)$$

Let us call the realization of this quantity R_1 . We then repeat the computation we have just described several more times to obtain the random realization

$$R_1, R_2, \dots, R_M$$

where M is a relatively large number.

The Monte Carlo approximation to the option price is then equal to

$$\frac{1}{M} \sum_{k=1}^M R_k.$$

A key insight here is that performing a risk-neutral simulation amounts to making an appropriate change in the parameters of the short rate.

11.3.6 SIMULATING INTEREST RATES IN ONE-FACTOR MODELS

Both the Vasicek and CIR models have the nice feature that the transition densities are known in closed form.⁵ In the case of the CIR model, the transition density is complicated, but it is possible to program a computer to take draws from this distribution.

We will discuss two ways to simulate interest-rates in these models:

⁵ For these models, the transition density determines the dynamics of the stochastic process. Therefore, knowing the transition density allows us to understand the role of the model parameters and to construct distributions for the outcomes of the stochastic process.

- Using exact transition densities
- Using discretized stochastic differential equations

The second method has the advantage that it can be implemented for a wide variety of interest rate dynamics and does not require explicit knowledge of the transition density. The main disadvantage of this method is that it generates approximate short-rate dynamics, and the discretization interval must be relatively small to obtain good results.

Simulation based on the use of discretized stochastic differential equations. Suppose we have a stochastic differential equation for the dynamics governing the short-rate process and a closed formula for bond prices. Let us suppose that the general stochastic differential equation is of the form

$$dr_t = \mu(r_t)dt + \sigma(r_t)dW_t$$

This equation may be discretized with discretization interval Δt as

$$r_{t+\Delta t} - r_t = \mu(r_t)\Delta t + \sigma(r_t)\sqrt{\Delta t}\varepsilon_{t+\Delta t}$$

where $\varepsilon_{t+\Delta t}$ is normally distributed with mean 0 and variance 1.

We will not spend time on the question of the rate of convergence of this approximation, but we do point out that there exist alternative discretization schemes that utilize more terms but have faster rates of convergence. If you were to employ such a scheme, the concepts we are discussing would remain the same except for the form of the discretization.

We may rearrange the discretization of the short rate as

$$r_{t+\Delta t} = r_t + \mu(r_t)\Delta t + \sigma(r_t)\sqrt{\Delta t}\varepsilon_{t+\Delta t}$$

If we can program our computer to generate draws from a normal distribution with mean 0 and variance 1, then we can use this equation to recursively generate a particular simulated trajectory for the short rate, starting from an initial value of r_0 .

As noted earlier, we will use the alternate notation $h = \Delta t$ to simplify expressions. For each trajectory in the simulation, we will thus have a sequence of values for the short-rate process

$$r_h, r_{2h}, r_{3h}, r_{4h}, \dots$$

out to some finite point in time representing the end of the simulation horizon.

This is exactly what we would obtain using the methodology of the previous section, the main difference being that here we do not require a knowledge of the exact transition density.

We may transform the simulated observations of the short rate into simulated term structures at each time along the trajectory, using the bond price function. Specifically, we obtain the simulated term structure

$$\{P(r_h, h, h+s) | s > 0\}, \{P(r_{2h}, 2h, 2h+s) | s > 0\}, \dots$$

Again, because we have a closed formula for bond prices, it is a simple matter to compute the term structure at each point along the trajectory, once we have simulated the short-rate evolution.

Let us briefly discuss the concrete cases of the Vasicek and the CIR models.

Implementation for the Vasicek model. The short-rate process for the Vasicek model is

$$dr_t = \alpha(\gamma - r_t)dt + \eta dW_t$$

This is then discretized as

$$r_{t+h} = r_t + \alpha(\gamma - r_t)h + \eta\sqrt{h}\varepsilon_{t+h}$$

where ε_{t+h} is normally distributed with mean 0 and variance 1.

As we have described, one now computes a sequence of realizations of the short rate for each trajectory, which results in the data

$$r_h, r_{2h}, r_{3h}, r_{4h}, \dots$$

This is then transformed into the sequence of term structures (in other words, yield curves) along the trajectory, using the bond price function. The sequence of term structures is

$$\{\exp(-r_h F(s) - G(s)) \mid s > 0\}, \{\exp(-r_{2h} F(s) - G(s)) \mid s > 0\}, \dots$$

Implementation for the CIR model. As you would anticipate, the implementation of the CIR model proceeds in much the same way as what we have just described for the Vasicek model. The one issue that is somewhat beyond the scope of our discussion is that the CIR model has a nonnegativity property that requires an additional adjustment when performing the discretization.

The short-rate process for the CIR model is

$$dr_t = \kappa(\theta - r_t)dt + \sigma\sqrt{r_t}dW_t$$

This is then discretized as

$$r_{t+h} = r_t + \kappa(\theta - r_t)h + \sigma\sqrt{r_t}\sqrt{h}\varepsilon_{t+h}$$

where ε_{t+h} is normally distributed with mean 0 and variance 1.

We now see the potential problem that can arise in implementing this discretization. Indeed, ε_{t+h} is normally distributed, so it can, in theory, take on a large negative value. If this were to happen, then r_{t+h} could take on a negative value. At the next iteration, the formula for r_{t+2h} would be

$$r_{t+2h} = r_t + \kappa(\theta - r_{t+h})h + \sigma\sqrt{r_{t+h}}\sqrt{h}\varepsilon_{t+2h}$$

and the algorithm would bomb, because we would be required to take the square root of a negative number.

There are two popular solutions to this problem. The first is to modify the discretization formula to

$$r_{t+h} = r_t + \kappa(\theta - r_t)h + \sigma\sqrt{\max(0, r_t)}\sqrt{h}\varepsilon_{t+h}$$

The second is to modify the discretization formula to

$$r_{t+h} = \max(0, r_t + \kappa(\theta - r_t)h + \sigma\sqrt{r_t}\sqrt{h}\varepsilon_{t+h})$$

The first step in simulating a trajectory for the evolution of the term structure is to compute a sequence of realizations of the short rate along a trajectory. We thus obtain the data

$$r_h, r_{2h}, r_{3h}, r_{4h}, \dots$$

This is then transformed into the sequence of term structures (in other words, yield curves) along the trajectory, using the bond price function. The sequence of term structures for the CIR model is

$$\{A(s) \exp(-r_h B(s)) \mid s > 0\}, \{A(s) \exp(-r_{2h} B(s)) \mid s > 0\}, \dots$$

11.3.7 MAXIMUM-LIKELIHOOD ESTIMATION FOR THE ONE-FACTOR VASICEK MODEL

The dynamics of the Vasicek model are based on the unobservable quantities (also called state variables) $\{r_t : t \geq 0\}$. The distribution of interest rates for the model is determined by the dynamics of the state variable. To estimate the model, we need to construct a metric that measures the fit of the model to the data used to calibrate it. Maximum-likelihood estimation (MLE) is one choice of metric. To apply MLE, we need to relate the distribution of interest rates implied by the model assumptions to the data (i.e., interest rates), so that we have a metric that can be optimized.

In theory, if we had N observations of the short rate, we could use these observations in the likelihood function

$$\mathcal{L} = f(r_1, r_2, \dots, r_N)$$

and find the usual MLE parameter estimates.

The main issue we must resolve is that we cannot observe these values, and we must compute the likelihood function from the data we can observe.

Fortunately, there is another way to approach the problem, which is based on information we can observe. We know from (VAS-PRICE) that there is a one-to-one correspondence between a time series of yield rates on a bond of a particular maturity and a time series of short rates.

Fix a time interval $h > 0$. We know that

$$P(r, t, t + h) = \exp(-rF(h) - G(h))$$

We may turn the algebra around and write

$$(EST) \quad r = \frac{-\ln P(r, t, t + h) - G(h)}{F(h)}$$

If we let $y(t, h)$ denote the yield to maturity on an h -year discount bond at time t , we can also express what we've just derived as

$$(EST-YIELD) \quad r = \frac{y(t, h)h - G(h)}{F(h)}$$

Since both $F(h)$ and $G(h)$ are constant for a fixed value of h , we have found that the unobservable short rates are linear functions of observable data.

Owing to the Markov property of the short-rate process, the likelihood function can be expressed as a product of conditional distributions as follows:⁶

⁶

If we include the current observation value r_0 , then the likelihood would change to $\mathcal{L} = f(r_0) \prod_{t=1}^N f(r_t | r_{t-1})$. The term $f(r_0)$ is the unconditional distribution of r_0 and in practice it is usually replaced with $f(\hat{r}_0)$ where \hat{r}_0 is the current observation of r_0 .

$$\mathcal{L} = f(r_1, r_2, \dots, r_N) = \prod_{t=1}^N f(r_t | r_{t-1}).$$

As we shall discuss in a moment, the conditional densities $f(r_t | r_{t-1})$ are known in both explicit and approximate form.

The key to completing the formulation of the estimation problem is to use (EST-YIELD) to establish the likelihood function.

Suppose that the random variable X has the density function f_X and that X is related to a second random variable Y through the linear equation

$$X = aY + b, \quad \text{with } a > 0.$$

By elementary algebra

$$\begin{aligned} P(Y \leq y) &= P\left(\frac{X - b}{a} \leq y\right) \\ &= P(X \leq ay + b) \end{aligned}$$

Therefore, we find that

$$f_Y(y) = f_X(ay + b)a$$

Applying this to the conditional density $f(r_t | r_{t-1})$ based on (EST-YIELD) yields the following formula for the likelihood function:

$$(LIKE) \quad \mathcal{L} = \prod_{t=1}^N f\left(\frac{-y(t, h)h - G(h)}{F(h)} \mid \frac{-y(t-1, h)h - G(h)}{F(h)}\right) \left(\frac{h}{F(h)}\right)$$

We see that the likelihood function is based on a known data set. We may then maximize this likelihood function if we know the conditional density function $f(\cdot | \cdot)$ in explicit form. There are two ways to obtain a conditional density function. The first is to use stochastic calculus to obtain the exact conditional density.

The second is to discretize the stochastic differential equation for the short rate and use this as a basis for an approximate conditional density. The approximate technique has the advantage that it will allow us to estimate models for which closed-form exact transition densities are not available. It has the disadvantage that the discretization process is known to introduce a bias in the estimation process. However, as the discretization interval tends to 0, this bias vanishes.

The approximate conditional density function. The approximate conditional density function is obtained by discretizing the stochastic differential equation for the short rate at a uniformly fixed discretization length.

Let us fix a uniform discretization length, which we will denote by Δt . The continuous-time dynamics of the short-rate are

$$dr_t = \alpha(\gamma - r_t)dt + \eta dW_t.$$

When we discretize this equation at discretization length Δt , we obtain the discrete time equation

$$r_{t+\Delta t} - r_t = \alpha(\gamma - r_t)\Delta t + \eta \sqrt{\Delta t} \varepsilon_{t+\Delta t}$$

where $\varepsilon_{t+\Delta t}$ is normally distributed with mean 0 and variance 1. We may rearrange this equation as

$$r_{t+\Delta t} = r_t + \alpha(\gamma - r_t)\Delta t + \eta\sqrt{\Delta t}\varepsilon_{t+\Delta t}$$

and again as

$$r_{t+\Delta t} = \alpha\gamma\Delta t + (1 - \alpha\Delta t)r_t + \eta\sqrt{\Delta t}\varepsilon_{t+\Delta t}.$$

Therefore, we see that conditional on r_t , the random variable $r_{t+\Delta t}$ has an approximate normal distribution with

$$\text{conditional mean} = \alpha\gamma\Delta t + (1 - \alpha\Delta t)r_t$$

$$\text{conditional variance} = \eta^2\Delta t.$$

We thus have an approximate conditional density for use in the likelihood function

$$f(r_{t+\Delta t} | r_t) = \frac{1}{\sqrt{2\pi\Delta t}\eta} \exp\left(-\frac{(r_{t+\Delta t} - [\alpha\gamma\Delta t + (1 - \alpha\Delta t)r_t])^2}{2\eta^2\Delta t}\right)$$

For obtaining maximum-likelihood estimates of the model parameters, it remains to numerically maximize the approximate likelihood function as a function of the four parameters, α , γ , η , and λ . This can be done using the built-in numerical functions in Excel.

The exact conditional density. We are interested in the transition density $f(r_s | r_t)$ for $s > t$. By using the method of integrating factors, we can solve the stochastic differential equation for the short rate and find that

$$r_s = e^{-\alpha(s-t)}r_t + \int_t^s e^{-\alpha(s-u)}\alpha\gamma du + \int_t^s e^{-\alpha(s-u)}\eta dW_u.$$

This implies that conditional on r_t , the random variable r_s is normally distributed.

We may then compute the conditional mean and variance using stochastic calculus as

$$E[r_s | r_t] = e^{-\alpha(s-t)}r_t + \gamma(1 - e^{-\alpha(s-t)})$$

and

$$\text{Var}[r_s | r_t] = \frac{\eta^2}{2\alpha}(1 - e^{-2\alpha(s-t)}).$$

We thus know the exact conditional distributions for use in the likelihood function:

$$f(r_s | r_t) = \frac{1}{\sqrt{2\pi\text{Var}[r_s | r_t]}} \exp\left(-\frac{(r_s - E[r_s | r_t])^2}{2\text{Var}[r_s | r_t]}\right)$$

Although the market price of risk parameter λA does not appear in the conditional density function, it does appear in (LIKE) because it appears in the function $G(h)$. (Recall that G is defined in Section 11.3.2.)

For obtaining maximum-likelihood estimates of the model parameters, it remains to numerically maximize the likelihood function as a function of the four parameters, α , γ , η and λ . One way to carry out this calculation is by using the built-in functions in Excel.

11.4 MORE ADVANCED MODELS FOR THE TERM STRUCTURE

The primary advantage of one-factor models is their simplicity. The most significant disadvantage of these models is that they cannot provide realistic models of the term structure.

The next step beyond a one-factor model is to introduce an additional factor, resulting in a two-factor model. Two-factor models are more realistic, but they provide this at a cost in computational complexity and estimation difficulties. We will discuss the most basic aspects of the two-factor Gaussian model.

One can move away from the most basic continuous-time affine models to a range of models classes such as these:

- Quadratic interest rate models
- Shadow interest rate models
- Threshold interest rate models
- Regime-switching models
- Macro finance models (such as the dynamic Nelson Siegel model)

Quadratic interest rate models can be single- or multifactor and have efficient procedures for computing prices and estimation. The quadratic class has some advantages in fitting swaption prices, and some development of their adaptation to low-interest-rate environments has been undertaken. *Shadow interest rate models* goes back to an idea put forward by Fischer Black and show some promise as a model class for the low-rate environments that have been prevalent post financial crisis. *Threshold interest rate models* relate the dynamics of interest rates to the level of the interest rate. Threshold models are computationally demanding. *Regime-switching models* relate the dynamics of interest rates to an unobservable latent process that is usually assumed to follow a Markov chain. The latent process serves to define the type of environment that the economy is currently in—for example, a low-volatility or a high-volatility regime. Regime-switching models have been extensively researched, and a range of effective techniques are available for computing prices and calibration. *Macro finance models* attempt to relate interest rate dynamics to important macroeconomic variables like GDP and inflation. The reader is referred to the annotated bibliography for some papers that cover these topics.

We now discuss details of some of the more advanced model classes.

11.4.1 A SIMPLE TWO-FACTOR GAUSSIAN MODEL

The term structure of interest rates is assumed to be driven by two factors, which we will call x and y . These are latent factors and do not necessarily directly correspond to observable economic variables. The dynamics of these two factors are assumed to follow the mean reverting Gaussian processes:

$$\begin{aligned} dx_t &= \kappa[\theta - x_t]dt + \sigma dW_t^{(1)} \\ dy_t &= K[\Theta - y_t]dt + \Sigma dW_t^{(2)}. \end{aligned}$$

It is assumed that these Brownian motions (driving noise) are independent. However, this assumption can be relaxed.

The short rate is denoted by r_t and is assumed to have the additive factor structure

$$r_t = x_t + y_t.$$

The market price of risk process for the factor x is assumed to be constant and equal to λ , and the market price of risk process for the factor y is assumed to be constant and equal to Λ .

The price of a discount bond at time t that matures for a unit at time s ($s > t$), when the factors are equal to x_t and y_t at time t , is denoted by

$$P(x_t, y_t, t, s).$$

Discount bond prices are determined by the usual risk-neutral expectation formula as:

$$(Bonds) \quad P(x_t, y_t, t, s) = E_t^{\mathbb{Q}} \left[\exp \left(- \int_t^s r_u du \right) \right].$$

Because the parameters in this model are time homogeneous, the bond price function will depend on t and s only through the difference $r \equiv s - t$. We may therefore economize in the notation we use to write our bond price as

$$P(x_t, y_t, \tau) \equiv P(x_t, y_t, t, s)$$

where τ is the time to maturity of the discount bond.

One can compute (Bonds) in closed form for this model.

Before we summarize these formulas, we require some notation. Note that the factor parameters appear in these expressions.

$$\begin{aligned} F_1(s) &= \frac{1 - e^{-\kappa s}}{\kappa} \quad \text{and} \quad F_2(s) = \frac{1 - e^{-Ks}}{K}. \\ D_1 &= \theta - \frac{\lambda\sigma}{\kappa} - \frac{\sigma^2}{2\kappa^2} \quad \text{and} \quad D_2 = \Theta - \frac{\Lambda\Sigma}{K} - \frac{\Sigma^2}{2K^2} \\ P_1(x, \tau) &= \exp \left(-x F_1(\tau) - \left[\tau D_1 - F_1(\tau) D_1 + F_1^2(\tau) \frac{\sigma^2}{4\kappa} \right] \right) \\ P_2(y, \tau) &= \exp \left(-y F_2(\tau) - \left[\tau D_2 - F_2(\tau) D_2 + F_2^2(\tau) \frac{\Sigma^2}{4K} \right] \right) \end{aligned}$$

The bond price function is given by the product

$$(BOND) \quad P(x, y, \tau) = P_1(x, \tau) P_2(y, \tau).$$

Simulating the term structure for the two-factor model. The simulation of the evolution of the term structure for this model now depends on the simulated two-dimensional trajectory of the two-dimensional process:

$$(x_t, y_t).$$

Providing that one has generated a trajectory of the state variables,

$$(TRAJ) \quad (x_h, y_h), (x_{2h}, y_{2h}), (x_{3h}, y_{3h}), \dots$$

one can then compute the term structure at each of these nodes as

$$\{P(x_h, y_h, s) | s > 0\}, \{P(x_{2h}, y_{2h}, s) | s > 0\}, \{P(x_{3h}, y_{3h}, s) | s > 0\}, \dots$$

This is in complete analogy to what we described for the one-factor case. The main difference is that we must simulate the trajectory of each factor to obtain the trajectory of the two state variables (TRAJ).

11.4.2 AFFINE MODELS

Affine models are a rich class of term structure models that was studied in its generality and popularized by the work of Duffie and Kan (1996). To this class belong many of the models developed in the pioneering works in the 1970s and '80s, including the Vasicek model (Vasicek 1977), the Cox, Ingersoll, and Ross (CIR) model (Cox et al. 1985a; 1985b) and other models. The first models were described by only one random process, but later on, these models were generalized to higher dimensions. Affine models assume that the dynamics of the term structure is driven by state variables, or factors. These factors are described by stochastic processes, and they represent the source of uncertainty of the models. They might be considered unobservable variables or nodes on the term structure, but in most applications, they are unobservable or latent variables, and we are going to describe in detail this case.

Affine models assume that the short rate, r_t , is a linear function of the state variables, X_t :

$$r_t = \delta_0 + \delta_1^T X_t$$

where $\delta_0 \in \mathcal{D} = \mathbb{R}$ and $\delta_1 \in \mathcal{D} = \mathbb{R}^d$, and that the state variable vector $X_t \in \mathcal{D} = \mathbb{R}_+^m \times \mathbb{R}^n$, $d = m + n$ follows an affine diffusion under the risk-neutral measure

$$dX_t = (b - \beta X_t)dt + \sigma(X_t)dW_t$$

with a given vector b and matrices β and σ . The instantaneous covariance matrix is

$$\sigma(x)\sigma(x)^T = a + \sum_{i=1}^m \alpha_i x_i$$

for matrices a and α_i .

In this setting, the state vector X_t is represented as a vector of m CIR processes and n Vasicek processes.⁷ We use the index sets $I = \{1, \dots, m\}$ and $J = \{m + 1, \dots, d\}$, and accordingly for a vector $u \in \mathbb{R}^d$, we write $u = (u_I, u_J)$. A diagonal matrix with u along the diagonal is denoted $\text{Diag}(u)$. The i th row of a matrix M we denote M_i . We note that a Feller factor x_i may be renormalized in such a way that $\alpha_{i,i} = 1$, and a Vasicek factor x_j may be renormalized in such a way that $b_j = 0$.

Using a no-arbitrage argument, one can prove that the bond prices are given by

$$P(t, T) = e^{\phi(T-t) + \psi(T-t)^T X_t}$$

(see Filipovic and Mayerhofer 2009, Theorem 4.1) with auxiliary functions that satisfy the following equations:

$$(4.1) \quad \dot{\phi} = -\delta_0 + b\psi + \frac{1}{2}\psi^T a\psi \quad \phi(0) = 0$$

$$(4.2) \quad \dot{\psi}_i = -\delta_{X,i} - \beta_i^T \psi + \frac{1}{2}\psi^T \alpha_i \psi, \quad i \in I \quad \psi_i(0) = 0$$

$$(4.3) \quad \dot{\psi}_j = -\beta_{j,j}^T \psi_j - \delta_{X,j} \quad \psi_j(0) = 0_j.$$

It can be shown that under certain regularity conditions on the coefficients, the solution of the partial differential equations exists and is unique. Moreover, under more restrictive assumptions on the dynamics of the state variables, the partial differential equations have a closed-form solution. One can see from the equations that in the affine framework, the yields are affine functions of the state variables, and more generally, all the quantities and prices of contingent claims depend on the state variables.

11.4.3 MARKET PRICE OF RISK

To completely specify the model, we have to provide a market price of risk, which gives the change of measure between the risk-neutral and the physical specifications of the model. Some conditions have to be satisfied in order to have an acceptable change of measure, but very often, more structure is imposed on the market price of risk in order to have affine dynamics under both measures. It turned out that the most general change of measure, introduced in Cheridito et al. (2007) and called extended affine market price of risk Λ_t , is given by

$$\Lambda_t = \sigma(X_t)^{-1}(\lambda_0 + \lambda_1 X_t)$$

⁷ CIR processes provide for nonnegative factors with noncentral chi-squared distributions, and Vasicek processes provide for factors with normal distributions. A combination of these factors provides some flexibility in the structure of the model.

Estimation. A very common technique used for the estimation of the model parameters is the maximum-likelihood estimation. Since in this class of models, the state variables are unobservable, a Kalman filter is typically used to extract the latent state variables from the historical yield series. There is an extensive literature on this. The interested reader can consult the following references:

- Singleton, Kenneth J. 2006. *Empirical Dynamic Asset Pricing*. Princeton, NJ: Princeton University Press.
- Duan, Jin-Chuan, and Jean-Guy Simonato. 1999. Estimating and Testing Exponential-Affine Term Structure Models by Kalman Filter. *Review of Quantitative Finance and Accounting* 13: 111–135.
- Chen, Ren-Raw, and Louis Scott. Maximum Likelihood Estimation for a Multifactor Equilibrium Model of the Term Structure of Interest Rates. *Journal of Fixed Income* 3(3): 14–31.

Fit of the initial term structure. Even if the model has been parameterized by using the historical time series, it is almost impossible that it can exactly fit an initial term structure (for instance, a zero-coupon curve or a swap curve). This drawback can be eliminated by using an exogenous extension of the models that preserves their analytical tractability. In other words, the short rate can be extended to be

$$r_t = l_t + \delta_1^T X_t$$

where $l_t : \mathbb{R}_+ \rightarrow \mathbb{R}$ is a deterministic piecewise function whose elements can be used to exactly fit the initial term structure. This method has been extensively described by Brigo and Mercurio (1998; 2001).

Pros and cons. The affine class of models has the following advantages:

- Under certain assumptions on the diffusion coefficients of the state variables, one can find closed bond price formulas and semiclosed derivative price formulas for vanilla and nonvanilla derivatives.
- This class of model can be used both under the risk-neutral and the real-world measures, because it has a realistic behavior under the physical measure and an accurate pricing under the risk-neutral measure.
- The affine structure allows well-defined estimation and calibration procedures, and the parameters can be efficiently estimated.
- The affine structure allows rich instantaneous volatility correlation structure of the state variables, which allows realistic correlations between yields within and across economies.

Its drawbacks are as follows:

- The models are linear, but in real data, nonlinearities can play an important role.
- The affine structure imposed some restrictions on the nature of some correlation behavior across yield tenors.

11.4.4 QUADRATIC MODELS

Quadratic models are a class of term structure models in which the short rate is a quadratic function of the state variables. This class of models has been studied in its generality by Ahn, Dittmar, and Gallant (1996) and encompasses other models previously developed by Longstaff (1989), Beaglehole and Tenney (1991; 1992), Constantinides (1992), Karoui, Myneni and Viswanathan (1992) and Ahn (1995).

Dynamics. In the quadratic term structure models, the short rate is a quadratic function of the state variables

$$r_t = X_t^T \gamma X_t + \delta_X^T X_t + \delta_0$$

where $\delta_0 \in \mathcal{D} = \mathbb{R}$, $\delta_X \in \mathcal{D} = \mathbb{R}^d$, $\gamma \in \mathcal{D} = \mathbb{R}^d \times \mathbb{R}^d$, and the state variable vector $X_t \in \mathcal{D} = \mathbb{R}^d$ follows a multivariate Gaussian process with mean reverting properties

$$dX_t = (b - \beta X_t)dt + \sigma dW_t$$

where $b \in \mathcal{D} = \mathbb{R}^d$ and $\beta, \sigma \in \mathcal{D} = \mathbb{R}^d \times \mathbb{R}^d$.

Moreover, it is assumed that β is diagonalizable with positive real components of eigenvalues, and dW_t is a d -dimensional vector of standard and independent Wiener processes.

To ensure positive interest rates, we require that γ be symmetric, positive semidefinite and that

$$(4.4) \quad 4\delta_0 \geq \delta_X^T \gamma^{-1} \delta_X.$$

The bond pricing formula is given by

$$P(t, T) = e^{X_t^T A(T-t) X_t + B(T-t)^T X_t + C(T-t)}$$

with auxiliary functions that satisfy

$$\begin{aligned} \dot{A} &= -\gamma - A\beta - \beta^T A - 2A\Sigma A, & A(0) &= 0 \\ \dot{B} &= -\delta_X + 2Ab - \beta^T B + 2A\Sigma B, & B(0) &= 0 \\ \dot{C} &= -\delta_0 + B^T b + \text{tr}(A\Sigma) - \frac{B^T \Sigma B}{2}, & C(0) &= 0 \end{aligned}$$

where $\Sigma = \sigma(X_t) \sigma^T$.

It can be shown that under certain regularity conditions on the coefficients, the solution of the partial differential equations exists and it is unique.

Market price of risk. To completely specify the model, we have to provide a market price of risk, which gives the change of measure between the risk-neutral and the physical specifications of the model. The most general change of measure, such that one has a multivariate Gaussian process under both measures, is given by

$$\Lambda_t = \sigma(X_t)^{-1}(\lambda_0 + \lambda_1 X_t)$$

where $\lambda_0 \in \mathcal{D} = \mathbb{R}^d$ and $\lambda_1 \in \mathcal{D} = \mathbb{R}^d \times \mathbb{R}^d$.

11.5.5 ESTIMATION

The estimation of this class of models can be done by using an extended Kalman filter, as in Claessens and Pennacchi (1996), or the efficient method of moments (EMM) of Gallant and Tauchen (1996).

One of the differences of quadratic models with respect to the affine models is that there is not a one-to-one correspondence between yields and state variables.

Pros and cons. The quadratic class of models has the following advantages:

- Flexibility in modeling heteroscedastic volatility and negative correlation among factors
- Potential to capture the nonlinearities in the data

It has the following drawbacks:

- Difficulties in fitting targets on both yields and returns
- A heavy and time-consuming estimation procedure

The dynamic Nelson-Siegel model. The dynamic Nelson-Siegel model has some desirable properties, and there has been a flurry of recent research into multiple aspects of the model, including the linking of the term structure to macroeconomic factors. A very useful general reference on the model is the book of Diebold & Rudebusch (2013).

The model is widely applied for forecasting interest rates and has several intuitively appealing features. In its most general form, the model is not arbitrage-free. However, it is possible to restrict this model to an arbitrage-free class.

The basis for implementation of this model is the equation

$$(4.5) \quad y_t(\tau) = \beta_{1t} + \beta_{2t} \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} \right) + \beta_{3t} \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right)$$

where λ is treated as a fixed parameter and β_{1t} , β_{2t} and β_{3t} are analogous to the role of state variables in affine modeling (Diebold and Rudebusch 2013, 26). When the model is fitted across panel data, β_{1t} , β_{2t} and β_{3t} form a time series of estimated parameters. For a discussion of the interpretation of the terms in this equation, see Diebold and Rudebusch (2013, 27).

Evidently, β_{1t} governs the level as it results in parallel shift, β_{2t} governs the slope, and β_{3t} governs the curvature. For this reason, (4.5) is often written as

$$(4.6) \quad y_t(\tau) = L_t + S_t \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} \right) + C_t \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right)$$

The idea behind the DNS model is quite simple. First, there is a “basis” consisting of three maturity-dependent functions, F_1 , F_2 and F_3 , which serve to determine the yield curve at each point in time. The basis functions are defined as follows:

$$(4.7) \quad F_1(\tau) = 1$$

$$(4.8) \quad F_2(\tau) = \frac{1 - e^{-\lambda\tau}}{\lambda\tau}$$

$$(4.9) \quad F_3(\tau) = \frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}$$

Second, the coordinates of the basis functions, β_{1t} , β_{2t} and β_{3t} , are taken as stochastic processes, so that a range of interesting and plausible yield curves are generated. This combination results in the DNS yield curve equation

$$(4.10) \quad y_t(\tau) = \beta_{1t} F_1(\tau) + \beta_{2t} F_2(\tau) + \beta_{3t} F_3(\tau).$$

Evidently, a different collection of basis functions could be used, but the primary consideration in the selection of the basis functions and the dynamics of the coordinate processes is maintaining the ability to capture a broad range of yield curve behavior.

There is a range of opinion on the nature of the parameter λ . Typically, λ is held fixed, based on some considerations relating to the most reasonable maturity at which to have maximum curvature effects. These considerations are outlined in Diebold and Rudebusch (2013, 346), and they fix this parameter at $\lambda = 0.0609$, since it is this value of the parameter for which the curvature factor achieves its maximum (at about 30 months’ maturity).

Direct calculation shows that

$$\frac{d}{d\tau} F_3(\tau) = \frac{\lambda e^{-\lambda t} (\lambda t) - 1 + \lambda e^{-\lambda t}}{(\lambda t)^2} + \lambda e^{-\lambda t}$$

Since there is no easily manipulated expression for the derivative of the third basis function with respect to maturity, the computation of the maximum level needs to be performed by numerical evaluation.

A most important element of understanding the DNS model is to determine the behavior of the data implied by coordinate processes β_{1t} , β_{2t} and β_{3t} . Key questions must be asked that enable us to determine what type of stochastic dynamics are needed to run the model. Some such key questions are the following:

- What is the path-wise behavior of the coordinate processes?
- Are the coordinate processes correlated?
- What regions of \mathbb{R}^3 do the coordinates need to occupy to capture historical yield curve behavior and to allow for extreme but plausible behavior?
- What types of stochastic processes can achieve the range of yield curve behavior we want to see with the Nelson-Siegel basis functions?

The data-implied coordinate processes are easy to compute if one fixes the value of λ . Figure 11.15 shows the properties of the coordinate processes for the United States.

Figure 11.15
TIME SERIES OF HISTORICAL FACTOR ESTIMATES FOR DNS MODEL



Note: $\beta_1 + \beta_2$ is the short rate.

Source: Conning, Inc.

Ordinary least squares (OLS) estimates for a fixed λ . We will compute the OLS estimates for the coordinate processes for a range of values of λ . We do not know, ex ante, how robust the OLS estimates will be to changes in λ , nor do we know how robust the OLS estimates will be to the number of points on the yield curve that are used in the OLS. We do know that the data window does not affect anything, since we are computing cross-sectional estimates. However, the data window does influence the stochastic dynamics used to model the coordinate processes.

To perform the OLS, we must specify the number of points on the yield curve that are used for the OLS. Obviously, we cannot use more points than we have data for, but we may not wish to use all cross-sectional data.

Let us denote the yield curve data maturity points that are to be used as

$$\tau_1, \tau_2, \dots, \tau_M$$

and the associated vector of points as τ . We are assuming that there are M yield curve data points included at each point in time.

We now form the matrix $M(\tau)$, which is defined as follows:

$$(4.11) \quad M(\tau) \stackrel{\text{def}}{=} \begin{bmatrix} F_1(\tau_1) & F_2(\tau_1) & F_3(\tau_1) \\ F_1(\tau_2) & F_2(\tau_2) & F_3(\tau_2) \\ \vdots & \vdots & \vdots \\ F_1(\tau_M) & F_2(\tau_M) & F_3(\tau_M) \end{bmatrix}$$

We also form the matrix of bond yield panel data:

$$(4.12) \quad \text{DATA}(\tau) \stackrel{\text{def}}{=} \begin{bmatrix} y_{t_1}(\tau_1) & y_{t_2}(\tau_1) & \cdots & y_{t_N}(\tau_1) \\ y_{t_1}(\tau_2) & y_{t_2}(\tau_2) & \cdots & y_{t_N}(\tau_2) \\ \vdots & \vdots & \vdots & \cdots \\ y_{t_1}(\tau_M) & y_{t_2}(\tau_M) & \cdots & y_{t_N}(\tau_M) \end{bmatrix}$$

where the columns of the data matrix consist of the time series of yield observations for the yield curve data maturity points that are to be used in the OLS estimation.

The estimation equation for the OLS is now

$$(4.13) \quad M(\tau) * B = \text{DATA}(\tau)$$

where the OLS estimated matrix

$$(4.14) \quad B \stackrel{\text{def}}{=} \begin{bmatrix} \hat{\beta}_{1,t_1} & \hat{\beta}_{1,t_2}(\tau_1) & \cdots & \hat{\beta}_{1,t_N}(\tau_1) \\ \hat{\beta}_{2,t_1} & \hat{\beta}_{2,t_2}(\tau_2) & \cdots & \hat{\beta}_{2,t_N}(\tau_2) \\ \hat{\beta}_{3,t_1} & \hat{\beta}_{3,t_2}(\tau_M) & \cdots & \hat{\beta}_{3,t_N}(\tau_N) \end{bmatrix}$$

contains the estimated panel of coordinate processes. The columns of B represent the coordinate estimates for the yield curve at a particular point in time. The rows of B are the data-implied time series for each coordinate process. These estimates are conditional on the value λ .

There is, of course, no guarantee that any given set of coordinates is a good fit to the data at that time. The ability of the DNS model to fit the data at any particular point in time is a function of the properties of the basis of maturity-dependent functions F_1, F_2, F_3 . In theory, one can augment the basis, as was done in the Svensson model, or change the basis functions in order to accommodate different yield curve shapes. Ideally, one is able to select a “good basis” that can capture a rich range of yield curve shapes, and one should find that the dynamics of the data-implied time series for the associated coordinate processes are amenable to a parsimonious stochastic model. If one finds that the data-implied time series for the associated coordinate processes are chaotic looking, then that means a reasonable model for the dynamics of the coordinate processes will not be able to capture the range of historical data that the OLS estimates were based on.

In summary, the art of the DNS modeling approach is to be able to find good basis functions and a rich but logical coordinate process dynamic. We must also use coordinate process dynamics that can be “easily” controlled and changed.

In MATLAB, the computation of B in equation (4.13) is performed as

$$B = M(\tau) \setminus \text{DATA}(\tau)$$

which is set up to solve the system, column by column, as

$$M(\tau) * B(:, j) = \text{DATA}(\tau)(:, j), j = 1, 2, \dots, t_N$$

and to write these column-by-column solutions back to B .

Comments on DNS. The DNS model provides an approachable and flexible framework for modeling interest rates. Strictly speaking, the DNS model is not arbitrage-free, but there are ways to amend the model to make it fit the arbitrage-free framework. The ability of the model to capture interest rate dynamics depends on the stochastic processes used to model the factor processes. An advantage of the DNS framework is that a wide range of factor dynamics can be introduced. Furthermore, the model is often integrated into a macro-finance framework, where the factor dynamics are directly linked to key macro variable dynamics. A drawback of the model is that it is not suitable for derivatives pricing and is therefore a tool for real-world applications.

11.5 CHAPTER REFERENCES

- Ahn, D. H. 1995. A Generalized Squared Autoregressive Intertemporal Term Structure Model. Working paper, University of North Carolina, Chapel Hill.
- Ahn, D. H., R. F. Dittmar and A. R. Gallant. 1996. Quadratic Term Structure Models: Theory and Evidence. *Review of Financial Studies* 15(1): 243–288.
- Beaglehole, D., and M. Tenney. 1991. General Solutions of Some Interest Rate–Contingent Claim Pricing Equations. *Journal of Fixed Income* 1: 69–83.
- . 1992. A Nonlinear Equilibrium Model of the Term Structure of Interest Rates: Corrections and Additions. *Journal of Financial Economics* 32: 345–454.
- Black, F., and P. Karasinski. 1991. Bond and Bond Option Pricing When Short Rates Are Lognormal. *Financial Analysts Journal* 47(4): 52–59.
- Brigo, D., and F. Mercurio. 1998. On Deterministic Shift Extensions of Short-Rate Models. Internal report, Banca IMI, Milan.
- . 2001. A Deterministic-Shift Extension of Analytically-Tractable and Time-Homogenous Short-Rate Models. *Finance and Stochastics* 5: 369–388.
- Chen, R. 1996. *Understanding and Managing Interest Rate Risks*. Singapore: World Scientific.
- Cheridito, P., D. Filipovic and R. L. Kimmel. 2007. Market Price of Risk Specifications for Affine Models: Theory and Evidence. *Journal of Financial Economics* 83: 123–170.
- Claessens, S., and G. Pennacchi. 1996. Estimating the Likelihood of Mexican Default From the Market Prices of Brady Bonds. *Journal of Financial and Quantitative Analysis* 31: 109–126.
- Constantinides, G. 1992. A Theory of the Nominal Term Structure of Interest Rates. *Review of Financial Studies* 5: 531–552.
- Cox, J. C., J. E. Ingersoll and S. A. Ross. 1985a. An Intertemporal General Equilibrium Model of Asset Prices. *Econometrica* 53: 363–384.
- . 1985b. A Theory of Term Structure of Interest Rates. *Econometrica* 53: 385–407.
- Dempster, M., and S. Pliska. 1997. *Mathematics of Derivative Securities*. New York: Cambridge University Press.
- Diebold, Francis X., and Glenn D. Rudebusch. 2013. *Yield Curve Modeling and Forecasting: The Dynamic Nelson-Siegel Approach*. Princeton, NJ: Princeton University Press.
- Dybvig, P. 1997. Bond and Bond Option Pricing Based on the Current Term Structure. In Dempster and Pliska (1997).
- Duffie, Darrell, and Rui Kan. 1996. A Yield-Factor Model of Interest Rates. *Mathematical Finance* 6: 379–406.
- Filipovic, D., and E. Mayerhofer. 2009. Affine Diffusion Processes: Theory and Applications. Working paper, Vienna Institute of Finance.
- Gallant, A. R., and G. Tauchen. 1996. Which Moments to Match. *Econometric Theory* 12: 657–681.
- Ho, T., and S. Lee. 2004. A Closed-Form Multifactor Binomial Interest Rate Model. *Journal of Fixed Income* 14(1): 8–16.
- Hull, J., and A. White. 1990. Pricing Interest-Rate-Derivative Securities. *Review of Financial Studies* 3(4): 573–592.
- Karoui, N. E., R. Myneni and R. Viswanathan. 1992. Arbitrage Pricing and Hedging of Interest Rates Claims with State Variables: I. Theory. Working paper, Université de Paris VI and Stanford University.
- Longstaff, F. 1989. A Nonlinear General Equilibrium Model of the Term Structure of Interest Rates. *Journal of Finance* 23: 1259–1282.

Lyuu, Y. 2004. *Financial Engineering and Computation: Principles, Mathematics, Algorithms*. New York: Cambridge University Press.

Tuckman, B. 2002. *Fixed Income Securities: Tools for Today's Markets*. 2nd ed. New York: John Wiley & Sons.

van der Hoek, J., and R. Elliott. 2006. *Binomial Models in Finance*. New York: Springer.

Vasicek, O. 1977. An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics* 5: 177–188.

11.6 APPENDIX: MATHEMATICAL DETAILS FOR A GENERAL SINGLE-FACTOR MODEL

The basis for single-factor models is the specification of a stochastic process for the short rate (also called the instantaneous riskless rate) as

$$dr = b(r, t)dt + a(r, t)dZ.$$

This equation is called a stochastic differential equation in which Z denotes a Brownian motion process (also called Wiener process). The term $b(r, t)$ is called the instantaneous drift of the short rate, and $a(r, t)$ is called the instantaneous standard deviation of the process. The term dZ is thought of as noise and represents random shocks to the level of the short rate. We can then obtain the prices of discount bonds and interest-rate-contingent claims as the solutions to certain partial differential equations.

11.6.1 MARKET PRICE OF RISK

It is assumed that all contingent claims can be expressed as a function of the short rate r , time t , and its maturity date T . Functionally, this means that the price of the claim can be expressed as

$$C(r, t, T).$$

As a convenience, we will refer to this generic claim as “the contingent claim C .” If the claim in question is an ordinary discount bond, it is customary to use the notation $P(r, t, T)$.

Using techniques from stochastic calculus (Itô’s lemma), we can show that C will satisfy the following stochastic differential equation:

$$\begin{aligned} \frac{dC(r, t, T)}{C(r, t, T)} &= \frac{1}{C} \left[C_t + b(r, t)C_r + \frac{1}{2}a^2(r, t)C_{rr} \right] dt \\ &\quad + \frac{1}{C}a(r, t)C_r dZ \end{aligned}$$

Let us introduce the notation

$$\alpha_C(r, t, T) = \frac{1}{C} \left[C_t + b(r, t)C_r + \frac{1}{2}a^2(r, t)C_{rr} \right]$$

and

$$\sigma_C(r, t, T) = \frac{1}{C}a(r, t)C_r.$$

We can then write our stochastic differential equation for C in the compact form

$$\frac{dC(r, t, T)}{C(r, t, T)} = \alpha_C(r, t, T)dt + \sigma_C(r, t, T)dZ.$$

Using sophisticated mathematics, we can show that in an arbitrage-free interest rate model, the ratio

$$\frac{\alpha_C(r, t, T) - r}{\sigma_C(r, t, T)}$$

cannot depend on T . In other words, this ratio is independent of which risky asset we form it from. We call this ratio the *market price of risk process* and denote it

$$\lambda(r, t) \equiv \frac{\alpha_C(r, t, T) - r}{\sigma_C(r, t, T)}.$$

We emphasize again that this ratio is the same for all risky assets in the model. The market price of risk is interpreted as the amount of excess return per unit of risk. It has the same feel and interpretation as the Sharpe ratio from portfolio theory.

11.6.2 DESCRIPTION OF THE BOND MARKET

The price $P(r, t, T)$ of a default-free zero-coupon bond is a function of the time t , the value of the short rate r , and the time $T \leq T^*$ of maturity of the bond. This is a fundamental assumption of the model. The particular form of the functional relationship is not specified; rather, we assume only that the price is a differentiable function of the three parameters.

By Itô's formula, the process $\{P(r_t, t, T) : 0 \leq t \leq T\}$ is also an Itô's process, and its differential equation is

$$dP = \left[\frac{1}{2} \sigma^2 \frac{\partial^2 P}{\partial r^2} + \mu \frac{\partial P}{\partial r} + \frac{\partial P}{\partial t} \right] dt + \sigma \frac{\partial P}{\partial r} dZ$$

where it is understood that the functions are evaluated at $(r, t) = (r_t, t)$. The drift and volatility components are written in this special form:

$$\mu_P(r, t, T) = \frac{1}{P} \left[\frac{1}{2} \sigma^2 \frac{\partial^2 P}{\partial r^2} + \mu \frac{\partial P}{\partial r} + \frac{\partial P}{\partial t} \right]$$

and

$$(4.15) \quad \sigma_P(r, t, T) = \frac{1}{P} \sigma \frac{\partial P}{\partial r}.$$

With this notational convention, the stochastic differential equation for the bond price is usually written in this equivalent form:

$$(4.16) \quad \frac{dP}{P} = \mu_P dt + \sigma_P dZ$$

The notation P , $P(t, T)$, and $P(t, r, T)$ all denote the price, but it is customary to suppress the reference to r , or even t and T , when no confusion can arise.

11.6.3 TRADING AND PORTFOLIOS

Continuous default-free borrowing and lending at the short rate is available to all traders. It is usually described in terms of an account, like a savings account. The account values from a process $\{B(t)\}$ satisfying $B(0) = 1$ and $dB(t) = r(t)B(t)dt$ for $t \geq 0$. The account has one dollar initially and grows by constantly earning and reinvesting the short rate. A trader who invests k dollars (buys the account) at time t_1 receives a right to $k/B(t_1)$ units of the account. At time $t_2 > t_1$, the trader can withdraw (sell units) with a value of $B(t_2)k/B(t_1)$.

The account value can be written explicitly as

$$B(t) = \exp\left(\int_0^t r(u)du\right)$$

so an investment of k at time t_1 accumulates to a value of

$$B(t_2)k/B(t_1) = k \exp\left(\int_{t_1}^{t_2} r(u)du\right)$$

Let S_1 and S_2 denote two traded securities in this market with differential equations

$$dS_i(t) = \mu_i(t)dt + \sigma_i(t)dZ(t)$$

for $i = 1, 2$. A portfolio of S_1 and S_2 is a pair of processes $(\Phi_1(t), \Phi_2(t))$, where $\Phi_i(t)$ denotes the number of units of the security i in the portfolio at time t . The number of units can be negative, indicating short selling (borrowing). The number of units need not be an integer; any real number is allowed for the number of units. The value of the portfolio at time t is

$$V(t) = \phi_1(t)S_1(t) + \phi_2(t)S_2(t).$$

The number of units in the portfolio has to be determined before the prices are revealed. This means that in setting a trading strategy, a trader may base the values of $(\Phi_1(t), \Phi_2(t))$ on information available for $s < t$ but not on information at time t . In addition, the trading strategy must satisfy the technical condition: For $i = 1, 2$:

$$\Pr\left[\int_0^T \phi_i(t)^2 \sigma_i(t)^2 dt < \infty\right] = 1 \quad \text{for any } T < \infty$$

A portfolio is self-financing provided that the portfolio's value changes over an interval $(t, t + dt)$ solely due to price changes, not due to changes in the number of units held. Mathematically, this is expressed as follows:

$$(4.17) \quad dV(t) = \phi_1(t)dS_1(t) + \phi_2(t)dS_2(t)$$

By substitution, we can find the Itô's representation of V :

$$(4.18) \quad dV(t) = \phi_1(t)[\mu_1(t)dt + \sigma_1(t)dZ] + \phi_2(t)[\mu_2(t)dt + \sigma_2(t)dZ]$$

$$(4.19) \quad = [\phi_1(t)\mu_1(t) + \phi_2(t)\mu_2(t)]dt + [\phi_1(t)\sigma_1(t) + \phi_2(t)\sigma_2(t)]dZ$$

Self-financing portfolios are a key ingredient in pricing and hedging. In applying the concept, we will usually arrange the strategy so that the portfolio's volatility term is zero and use the following additional important assumption.

11.6.4 NO-ARBITRAGE PRINCIPLE

An arbitrage opportunity is a self-financing portfolio with zero initial value, having at some future time no possibility of a negative value and a positive probability of a positive value. In terms of the notation defined above, an arbitrage opportunity based on the securities S_1 and S_2 is a portfolio $(\Phi_1(t), \Phi_2(t))$ for which the value process $V(t) = \Phi_1(t)S_1(t) + \Phi_2(t)S_2(t)$ satisfies the following conditions:

- $dV = \phi_1 dS_1 + \phi_2 dS_2$
- $V(0) = 0$
- There is a $T > 0$ for which

$$\Pr(V(T) < 0) = 0 \text{ and } \Pr(V(T) > 0) > 0.$$

The bond market is assumed to be arbitrage-free, meaning there are no arbitrage opportunities. This natural assumption has some strong implications for modeling. A common application goes like this: If a traded security (or portfolio of traded securities) with price $\{X_t\}$ has zero volatility, then its drift is $r_t X_t$, where r_t is the short rate. By zero volatility, we mean that, when written in its Itô representation,

$$dX = \mu_X(t)dt + \sigma_X(t)dZ,$$

the volatility term is identically zero with probability one:

$$\Pr[\sigma_X(t) = 0 \text{ for all } t] = 1$$

For any such security or portfolio, we must have $\mu_X(t) = r(t)X(t)$ for all t with probability one.

11.6.5 MARKET PRICE OF RISK

Consider two zero-coupon bonds denoted by $S_1(t) = P(r, t, T_1)$ and $S_2(t) = P(r, t, T_2)$ with $T_2 < T_1$. The zero-coupon bonds have differential equations $dS_i = \mu_i S_i dt + \sigma_i S_i dZ$. Make a portfolio with one unit of S_1 and $\phi(t)$ units of S_2 at time t , selected so that the resulting portfolio has zero drift:

$$V(t) = S_1(t) + \phi(t)S_2(t)$$

The differential equation of $V(t)$ is

$$\begin{aligned} dV(t) &= [\mu_1(t)S_1(t) + \phi(t)\mu_2(t)S_2(t)]dt \\ &\quad + [\sigma_1(t)S_1(t) + \phi(t)\sigma_2(t)S_2(t)]dZ \end{aligned}$$

For a diffusion coefficient of zero, $\phi(t) = -S_1(t)\sigma_1(t)/S_2(t)\sigma_2(t)$ for all t . Since there is no arbitrage and the portfolio has zero volatility, then its drift is $r(t)V(t)$. Substituting the values of the parameters and solving gives us the following relation:

$$\begin{aligned} r(t)V(t) &= \mu_1(t)S_1(t) + \phi(t)\mu_2(t)S_2(t) \\ r(t)[S_1(t) + \phi(t)S_2(t)] &= \mu_1(t)S_1(t) - S_1(t)\mu_2(t)\sigma_1(t)/\sigma_2(t) \\ r(t)[S_1(t) - S_1(t)\sigma_1(t)/\sigma_2(t)] &= \mu_1(t)S_1(t) - S_1(t)\mu_2(t)\sigma_1(t)/\sigma_2(t) \\ r(t)[\sigma_2(t) - \sigma_1(t)] &= \mu_1(t)\sigma_2(t) - \mu_2(t)\sigma_1(t)/\sigma_2(t) \\ r(t) &= \frac{\mu_1(t)\sigma_2(t) - \mu_2(t)\sigma_1(t)}{\sigma_2(t) - \sigma_1(t)} \end{aligned}$$

This can be rearranged to obtain

$$\frac{\mu_2(t) - r(t)}{\sigma_2(t)} = \frac{\mu_1(t) - r(t)}{\sigma_1(t)}$$

under our original assumption that these are different bonds.

Thus, the no-arbitrage principle forces this relation: The excess of a bond's drift over the default-free rate divided by its volatility is independent of the maturity of the bond. The common value is a characteristic of the market rather than the security. This quantity is called the market price of risk. Although it is not a price, the name is widely used.

So far, we do not have a formula for valuing zero-coupon bonds (or any other security) in terms of the short rate. The missing ingredient is the market price of risk. Therefore, in addition to our other assumptions, we assume that the market price of risk, denoted $\lambda(r, t)$, is known. Its relation to the parameters of the zero-coupon bond $P(r, t, T)$ is

$$(4.20) \quad \begin{aligned} \lambda(r, t) &= \frac{P(t, r, T)\mu_P(r, t, T) - rP(t, r, T)}{\sigma_P(r, t, T)P(t, r, T)} \\ &= \frac{\mu_P(r, t, T) - r}{\sigma_P(r, t, T)} \end{aligned}$$

where

$$\frac{dP(r, t, T)}{P(r, t, T)} = \mu_P(r, t, T)dt + \sigma_P(r, t, T)dZ$$

for all zero-coupon bond maturities T . There are important consequences of these assumptions. The equation becomes a partial differential equation with known coefficients when we (again) rewrite it, but now in terms of the bond price function and its derivatives:

$$\begin{aligned} \lambda \sigma_P &= \mu_P - r \\ \lambda P \sigma_P &= (\mu_P - r)P \\ \lambda \sigma \frac{\partial P}{\partial r} &= \frac{1}{2}\sigma^2 \frac{\partial^2 P}{\partial r^2} + \mu \frac{\partial P}{\partial r} + \frac{\partial P}{\partial t} - rP \end{aligned}$$

Therefore, we can determine zero-coupon bond prices in terms of the short rate drift, diffusion, and the market price of risk by solving this partial differential equation:

$$(4.21) \quad \frac{1}{2}\sigma^2 \frac{\partial^2 P}{\partial r^2} + (\mu - \lambda\sigma) \frac{\partial P}{\partial r} + \frac{\partial P}{\partial t} - rP = 0$$

for $0 < r < \infty$, $0 < t < T$, subject to the boundary conditions: $P(T, T) = 1$ and $r(0) = r_0$ is the observed value of the short rate at $t = 0$.

We also rewrite the bond's stochastic differential equation incorporating the market price of risk:

$$(4.22) \quad \begin{aligned} \frac{dP}{P} &= \mu_P dt + \sigma_P dZ \\ &= (r + \lambda \sigma_P)dt + \sigma_P dZ \end{aligned}$$

11.6.6 VALUATION OF INTEREST-RATE-CONTINGENT CLAIMS

We will now derive the partial differential equation (PDE) that the price of a general interest-rate-contingent claim must satisfy.

As we have shown, the instantaneous return and instantaneous standard deviation of the claim satisfies the restriction

$$\frac{\alpha_C(r, t, T) - r}{\sigma_C(r, t, T)} = \lambda(r, t),$$

for some market price of risk process. This can be rearranged as

$$\alpha_C(r, t, T) = r + \lambda(r, t)\sigma_C(r, t, T).$$

Substituting the full expressions for $\alpha_C(r, t, T)$ and $\sigma_C(r, t, T)$ in this equation results in the following equation:

$$\begin{aligned} \frac{1}{C} \left[C_t + b(r, t)C_r + \frac{1}{2}a^2(r, t)C_{rr} \right] \\ = r + \lambda(r, t)\frac{1}{C}a(r, t)C_r. \end{aligned}$$

Simplifying this and rearranging gives the PDE that is satisfied by a general interest-rate-contingent claim:

$$\begin{aligned} \frac{1}{2}a^2(r, t)C_{rr} + [b(r, t) - \lambda(r, t)a(r, t)]C_r - rC + C_t \\ = 0. \end{aligned}$$

Generally, the price of an interest-rate-contingent claim is determined as the solution to this equation with appropriate boundary conditions. The boundary conditions usually reflect the price of the claim at its maturity. For instance, if the claim is a discount bond, denoted P , maturing at time T , then the appropriate boundary condition is

$$P(r, T, T) = 1.$$

Similarly, if the claim is an interest rate cap, denoted C , with a notional amount of \$1,000 and a strike price of x that pays once at time T , then the appropriate boundary condition is

$$C(r, T, T) = 1,000 \max\{r - x, 0\}$$

It may be necessary to impose additional regularity conditions.

11.6.7 FEYNMANN-KAČ FORMULA

The price at time t is a function of t and the current value of the short rate $r_t = r$. The price of a zero-coupon bond is the expected discounted value of its future cash payment. In general, the price of a security that pays its owner a single payment of $H(r_T)$ at time T is also the expected discounted value of its cash payment. The expectation is taken over the paths described by the differential equation

$$dr_t = [\mu(r_t, t) - \lambda(r_t, t)\sigma(r_t, t)]dt + \sigma(r_t, t)dZ_t$$

subject to $r_t = r$. This follows from the Feynmann-Kač formula.

Let $\{X_t\}$ denote an Itô process with equation $X_t = \mu(X_t, t)dt + \sigma(X_t, t)dZ_t$. Let $H(x)$ be a given function for the states of X_T . Define the function:

$$V(x, t) = E[H(X_T) \exp(-\int_t^T X_u du) | X_t = x]$$

Then $V(x, t)$ satisfies the partial differential equation

$$\frac{\partial V(x, t)}{\partial t} + \mu(x, t)\frac{\partial V(x, t)}{\partial x} + \frac{1}{2}\sigma^2(x, t)\frac{\partial^2 V(x, t)}{\partial x^2} - xV(x, t) = 0$$

subject to $V(x, T) = H(x)$ for all x .

This is an intuitive justification for the formula. Let

$$Y_t = \exp(-\int_t^T X_u du) \text{ for } t \leq T.$$

The derivative of the path integral $-\int_t^T X_u du$ with respect to the lower limit t is $X_t dt$.

So by the Itô formula with

$$f(x, t) = V(x, t),$$

the differential of

$$Y_t = V(X_t, t).$$

Another application of the Itô formula with $f(x, t) = e^x$ shows that the differential of $Y_t = V(X_t, t)$ is

$$dY_t = \left[\frac{\partial V(X_t, t)}{\partial t} + \mu(X_t, t) \frac{\partial V(X_t, t)}{\partial x} + \frac{1}{2} \sigma^2(X_t, t) \frac{\partial^2 V(X_t, t)}{\partial x^2} \right] dt + \sigma(X_t, t) \frac{\partial V(X_t, t)}{\partial x} dZ_t.$$

Now equate the two conditional expectations. The first representation gives

$$\mathbb{E}[dY_t | X_t = x] = \mathbb{E}[X_t Y_t dt | X_t = x] = xdt \mathbb{E}[Y_t | X_t = x] = xV(x, t)dt.$$

The second gives

$$\begin{aligned} \mathbb{E}[dY_t | X_t = x] &= \mathbb{E}\left[\left. \frac{\partial V(X_t, t)}{\partial t} + \mu(X_t, t) \frac{\partial V(X_t, t)}{\partial x} + \frac{1}{2} \sigma^2(X_t, t) \frac{\partial^2 V(X_t, t)}{\partial x^2} \right| X_t = x\right] dt \\ &\quad + \mathbb{E}\left[\left. \sigma(X_t, t) \frac{\partial V(X_t, t)}{\partial x} dZ_t \right| X_t = x\right] \\ &= \left[\frac{\partial V(x, t)}{\partial t} + \mu(x, t) \frac{\partial V(x, t)}{\partial x} + \frac{1}{2} \sigma^2(x, t) \frac{\partial^2 V(x, t)}{\partial x^2} \right] dt + \sigma(x, t) \frac{\partial V(x, t)}{\partial x} \mathbb{E}[dZ_t] \\ &= \left[\frac{\partial V(x, t)}{\partial t} + \mu(x, t) \frac{\partial V(x, t)}{\partial x} + \frac{1}{2} \sigma^2(x, t) \frac{\partial^2 V(x, t)}{\partial x^2} \right] dt \end{aligned}$$

Equating the two expressions yields the Feynmann-Kač formula. The boundary condition follows from the fact that $Y_T = H(x)$.

This is applied to bond markets with the short-rate process $\{r_t\}$ in place of $\{X_t\}$. The physical short-rate process is

$$dr_t = \mu(r_t, t)dt + \sigma(r_t, t)dZ_t.$$

The price of a short-rate derivative at time t with current short rate r is denoted $V(r, t)$. It pays its owner a single payment of $H(r_t)$ at time T . As before, it satisfies the partial differential equation

$$\frac{\partial V(r, t)}{\partial t} + [\mu(r, t) - \lambda(r_t, t)\sigma(r_t, t)] \frac{\partial V(r, t)}{\partial r} + \frac{1}{2} \sigma^2(r, t) \frac{\partial^2 V(r, t)}{\partial r^2} - rV(r, t) = 0$$

subject to $V(r, T) = H(r)$ for all r .

By the Feynmann-Kač formula the function

$$V^*(r, t) = \mathbb{E}^* \left[H(r_T) \exp \left(- \int_t^T r_u du \right) \middle| r_t = r \right]$$

satisfies the same partial differential equation and boundary condition where the short-rate process is driven by

$$dr_t = [\mu(r_t, t) - \lambda(r_t, t)\sigma(r_t, t)]dt + \sigma(r_t, t)dZ_t^*.$$

Hence, $V^*(r, t) = V(r, t)$. We have two methods of numerically calculating the derivative price. We can solve the partial differential equation numerically, or we can calculate the expected discounted value by simulating the short-rate

process, being careful to use the drift $\mu - \lambda\sigma$, and evaluate the value $H(r_T) \exp\left(-\int_t^T r_u du\right)$ many times. Then, the average of the simulated values approximates the price.

11.6.8 TIME HOMOGENEITY AND THE τ NOTATION

When the parameters of the short-rate process do not depend on time, the price of interest-rate-contingent claims (including discount bond prices) can be expressed in the following form:

$$C(r, t, T) = C(r, T - t) \equiv C(r, \tau)$$

where $\tau = T - t$. By the chain rule of ordinary calculus

$$\begin{aligned} \frac{\partial C(r, T - t)}{\partial t} &= \frac{\partial C(r, T - t)}{\partial \tau} \frac{\partial(T - t)}{\partial t} \\ &= \frac{\partial C(r, T - t)}{\partial \tau} (-1). \end{aligned}$$

In other words, we find that

$$C_t = -C_\tau.$$

Consequently, when contingent claims have this functional property, the valuation PDE can be simplified by replacing C_t with $-C_\tau$.

In models for which the short-rate process is time homogeneous (which means the drift function $b(r, t)$ and instantaneous volatility of the short rate $a(r, t)$ do not depend on time t and are thus of the form $b(r)$ and $a(r)$ respectively), discount bond prices will be of the form

$$P(r, T - t) \equiv P(r, \tau).$$

As we have examined in the main body of this chapter, the Vasicek model

$$dr = \alpha(\gamma - r)dt + \eta dW$$

and the Cox-Ingersoll-Ross model

$$dr = \kappa(\theta - r)dt + \sigma\sqrt{r}dW$$

are two examples of such short-rate processes that lead to discount bond prices of the form $P(r, \tau)$.

Chapter 12: Corporate Bond Models

A corporate bond is a security with a fixed schedule of payments, generally consisting of a scheduled maturity date, and associated coupons and a redemption amount. The receipt of the scheduled payments is contingent upon the issuing name being able and willing to pay them as scheduled. Similar to an investor in default-free bonds, the corporate bond investor is subject to fluctuation in prices resulting from the general level of interest rates, but in contrast to default-free bonds, pricing will also contemplate the possibility that these payments may not be made as scheduled. There are also pricing implications resulting from changes in the potential for default. The possibility or belief that such an event will occur or has occurred may fluctuate with specific conditions of the issuer or class of issuers, so the impact on pricing may also fluctuate. The general idea of fluctuations in the price of a bond or a class of bonds is referred to as migration risk.

The corporate bond market is the universe of all existing corporate bonds, together with the institutional details necessary for the market to function. Corporate bond models address these further complications and institutional details regarding these classes of securities. The mathematics required to represent these complications can be intense, and this chapter presents an introduction to how models should go about reflecting this complexity.

Section 12.1 addresses specific considerations that corporate bond models need to address concerning corporate bonds and the corporate bond market. These include default and the potential for recovery from default; signals for the potential of default (including ratings) and migration of changes in those ratings; and credit spread dynamics.

Section 12.2 introduces some basic mechanisms of corporate bond models, including four classes of corporate bond models (structural or Merton models, reduced-form models, ratings-based models and hybrid models). The section also provides an overview of pricing mechanics.

Section 12.3 introduces structural or Merton models. These models are based on modeling the value of the firm that issued the debt to be priced. The value of the corporate debt is assumed equal to the firm value less the value of the equity.

Section 12.4 discusses reduced-form models. These models do not contemplate ratings changes during the life of the bond. They consider a risk-free term structure as a component, along with credit risk as an intensity process. Additional considerations that may be incorporated in the model are loss given default, arbitrage-free pricing, zero-recovery considerations and the role of risk premium.

Section 12.5 addresses ratings-based models, beginning with a discussion of a basic transition matrix to model changes in the spread of a corporate bond in response to changes in the underlying rating. It presents the mathematics behind the development of this transition matrix, along with considerations necessary to make this model tractable.

Section 12.6 presents additional considerations, including a brief discussion of validation and considerations regarding the uncertainty in corporate bond behavior introduced after 2008.

12.1 DEFINITION AND KEY CONSIDERATIONS OF THE CORPORATE BOND MARKET

By the phrase *corporate bond market*, we mean the universe of all existing corporate bonds together with the institutional details necessary for the market to function. These institutional details include redemptions, new issuance, trading protocols, call provisions, and resolution procedures for default events. There is no doubt that the corporate bond market is a complex market of bonds. It is therefore not surprising that seemingly innocuous concepts such as the “AA yield curve” turn out to be more involved than casual reflection may suggest. Generally, significant but reasonable assumptions must be made about what proxy for the corporate bond market a particular corporate bond model is attempting to capture. For example, it is common to assume that a corporate bond model is restricted to the modeling of noncallable industrial-issuer bonds. An effective ESG will need to have all of these elements if the regulatory aspects of corporate bonds are to be captured.

For our purposes, a corporate bond is a security with a fixed schedule of payments, generally consisting of a scheduled maturity date and associated coupons and a redemption amount.

What are the main risk factors for an investor who owns corporate bonds?

12.1.1 DEFAULT

Certainly, the bond may default. If the bond does default, will any further payments be received by the investor, thereby representing a partial recovery of capital? The receipt of the scheduled payments is contingent upon the issuing name being able and willing to pay them as scheduled. In contrast to default-free bonds (also referred to as treasury bonds), there is a possibility that these payment may not be made as scheduled. If the payments schedule is broken, default is said to have occurred. The probabilities of default will fluctuate with general economic conditions and firm- or industry-specific conditions. A wide range of default outcomes are possible, ranging from a total loss of all remaining scheduled payments to relatively small losses or an extension of the payment schedule. Either a given bond will survive to its maturity date without migrating to default, or it will migrate to default prior to its scheduled maturity date. The events that occur after a bond defaults are complicated and vary case by case. For modeling purposes, a recovery protocol is specified to consistently handle bonds that migrate to default prior to their scheduled maturity date.

12.1.2 RATINGS AS A CREDIT RISK SIGNAL

Another important risk factor for an investor that holds a bond is that the bond issuer's ability to meet its scheduled obligations might decline. The rational response to knowing for certain that such an event has occurred is to place a lower value on the bond's scheduled payments (i.e., future promises), resulting in a reduced price for the bond. The rational response to observing a highly reliable (but not totally reliable) signal that the bond issuer's ability to meet its scheduled obligations has declined is also a reduced price of the bond. In practice, the signal that most corporate bond investors receive is a rating downgrade.

Most corporate bonds of interest to insurers will be rated during their lives. These ratings are provided by an independent rating agency, and the bond's rating endeavors to capture the likelihood the bond will make its scheduled payments. The rating for an individual bond will generally fluctuate over time. If the rating is to have any practical value, one would expect that a deterioration in a bond's rating would be accompanied by a decline in its price, and that an improvement in a bond's rating would be accompanied by an increase in its price. To some extent, the success of a corporate bond model that includes rating dynamics lies in capturing this simple idea in a sufficiently robust fashion that permits a realistic but rich range of co-movements in ratings and price. For instance, during the financial crisis, ratings did not change even in the face of large spread widening. Ratings are also necessary for insurance companies to comply with risk-based capital requirements.

However, the act of a rating agency announcing to the world that the bond the investor holds is now less likely to make its payments than it was before should not affect the price of the bond unless the market believes that this signal is reliable. Recent experience shows that ratings have differential value. In a case where an investor does not regard a rating as credible, that investor might elect to continue to hold the bond after a reduced rating, and might even plan to hold the bond to maturity. Evidently, if that bond does not default, then the investor will receive all cash flows as promised, even though the investor had to endure the indignity of holding a bond that a rating agency had downgraded, perhaps more than once. Obviously, the opposite sort of situation may occur as well, whereby an investor holds a bond with a strong credit rating, but the bond nonetheless goes into default.

Changes in bond ratings do have real consequences for risk-based capital requirements and investment portfolio rules, even if default never happens.

ESGs typically adopt the position that a rating is a correlated signal of the creditworthiness of a bond. For modeling purposes, this means that there must be a meaningful relationship between a bond's rating and its price. There are two fundamental ways to incorporate ratings into a corporate bond model:

- Rating changes may be specifically embedded in the dynamics.
- Ratings may be implied by the movements in bond prices; we call these market-implied ratings.

12.1.3 MIGRATION RISK AND CREDIT SPREAD DYNAMICS

The general idea that a bond's rating will fluctuate and thereby induce fluctuations in the price of that bond is referred to as migration risk. The corporate bond investor is also subject to fluctuation in the general level of interest rates in the same way that an investor in default-free bonds (i.e., treasury bonds) is. Since changes in the general level of interest rates are not a risk that is specific to the corporate bond market, they are not a risk that defines the purpose of a corporate bond model. Instead, the power of a corporate bond model is distinguished by its ability to accurately capture credit spread dynamics. Credit spreads may be thought of as a broad measure of the additional yield over a comparable treasury security that an investor gets from holding a corporate bond. Credit spread dynamics are then the model of how this risk-return relationship varies through the simulation.

12.2 THE MECHANISMS OF CORPORATE BOND MODELING

Corporate bond modeling is considered part of the general area of credit risk modeling. A corporate bond model should generate prices, cash flows, and simulated dynamics for a realistic corporate bond universe. Ideally, a mechanism for attaching a rating to each bond also should be provided. The model must be calibrated to the appropriate historical benchmarks and validated against them.

We will focus on bonds without call provisions or other derivatives embedded in them.

Many approaches to credit risk modeling have been used. These include four broad classes of models:

1. Structural models, also referred to as Merton models
2. Reduced-form models, which have an actuarial feel and resemble survival distribution theory¹
3. Ratings-based models
4. Hybrid credit models, which may involve copulas, may not be arbitrage-free, and are often focused on important risk factors of a specific credit application

There are many approaches to the modeling of corporate bonds in arbitrage-free and more general modeling contexts. We are interested in continuous-time models, the primary focus of which is to model credit spreads and to capture certain stylized facts.

We adopt the simplified view of the corporate bond market as a universe of coupon-bearing bonds, each of which will have a rating at each point in time. The ratings that the bonds may be assigned are AAA, AA, A, BBB, HY and Def. Each bond is a vanilla coupon bond with no call features or other special provisions.

We adopt the recovery of market value (RMV) protocol when we provide technical model details below. Under the RMV protocol, the defaulted bond is assumed to be immediately retired at the time of default, and the bondholder receives a payment at the time of default equal to a fraction of the market value of the defaulted bond on the instant prior to its default. The fraction of the market value received is modeled by a loss given default (LGD) variable, which is often taken to be a constant. The concept of LGD is actuarial in nature and corresponds to the "severity" portion of the classical actuarial frequency-severity model. Protocols other than RMV may be used, each with its relative pros and cons. A popular alternative recovery protocol is recovery of face amount. We will use this protocol to present a heuristic overview of pricing in Section 12.2.1.

12.2.1 A HEURISTIC OVERVIEW OF CREDIT-PRICING MECHANICS

We consider the problem of pricing a defaultable bond at time t that is scheduled to make a series of payments up until time T .

We use $v(t; T)$ to denote the value at time t of one unit scheduled to be paid at time T , where $T > t$ and the scheduled payment is subject to credit risk.

¹ The seminal paper by Duffie and Singleton (1999) popularized these models, but aspects of this class of models predate that work.

We will denote the time that the bond defaults by the symbol r . The time of default may be thought of as the financial analogue of the time of death in life insurance mathematics. Unlike life insurance mathematics, where the ultimate time of death is usually considered, in credit risk pricing we are typically only concerned with whether default occurs before the scheduled maturity date of the bond. The event that the bond defaults prior to the scheduled maturity date of the bond corresponds to the set $\{r < T\}$.

Let us consider the pricing of a defaultable zero-coupon bond scheduled to mature for a unit amount at time T . For the moment, we assume that if default occurs, then all of the remaining scheduled payments are forfeited.² One might reason that the price of this bond can be computed as the present value of a unit amount to be delivered at time T , weighted by the probability that the bond has not defaulted by time T , i.e., the bond survives to time T .

Let $Q_0(r > T)$ denote the risk-neutral probability that the newly issued bond at time zero survives to maturity. Then we are reasoning that the price of the defaultable bond should look something like this:

$$(12.1) \quad v(0, T) = P(0, T)Q_0(\tau > T).$$

Note that Equation 12.1 says the spread to treasury³ on such a bond is equal to

$$(12.2) \quad -\frac{1}{T} \log(Q_0(\tau > T)).$$

The pricing expression in Equation 12.1 may be generalized to time $t > 0$ as follows.

Let $Q_t(r > T)$ denote the risk-neutral probability that the bond survives to maturity. Then we are reasoning that the price of the defaultable bond should look something like this:

$$(12.3) \quad v(t, T) = P(t, T)Q_t(\tau > T)$$

with associated credit spreads

$$(12.4) \quad -\frac{1}{T-t} \log(Q_t(\tau > T)).$$

Under certain conditions, Equation 12.3 is valid.

Evidently, under pricing formulas such as Equation 12.3, there is independence of the default-free term structure from the credit spread component.

It is plain from the expression in Equation 12.4 that the dynamics of the credit spreads depend on the dynamics of the risk-neutral default probabilities. If the risk-neutral default probabilities are deterministic, meaning that they depend only on the time and/or the remaining time to maturity, then credit spreads will be deterministic.⁴ This is unrealistic, since credit spreads fluctuate over time.

The pricing of defaultable coupon bonds then flows through by pricing the individual components of the coupon bond and adding the result.

12.3 STRUCTURAL/MERTON MODELS

Structural models are based on a model for the value of the firm that issued the debt to be priced. Some simplifying assumptions are needed to get the theory into a tractable form, but the intuitions the model provides are valuable.

² This assumption is referred to as “zero recovery.”

³ Measured as a continuously compounded interest rate. In general the spread over treasury for a T -year zero-coupon bond is equal to $-\frac{1}{T} \log(v(0, T)/P(0, T))$.

⁴ For example, when risk-neutral default probabilities are deterministic, then the spread to treasury for a 10-year zero-coupon bond will be a constant given by $-\frac{1}{10} \log(Q_t(\tau > t + 10))$.

Merton's original paper "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates" remains a very readable presentation and an important source from which to understand this model.

The ideas proceed broadly as follows: A stochastic process is defined for the value of the firm, generally referred to as firm value. One may also think of this as the value of the firm's assets. It is assumed that the value of the firm comprises two components: the value of the equity issued by the firm and the value of the debt issued by the firm. Assume that the debt issued by the firm matures for a face amount F . At the time the debt matures, the holders of the debt will receive the smaller of the redemption amount or the value of the firm.⁵ If the value of the firm is less than the amount the firm's debt matures for, then the firm is technically bankrupt, a default trigger.

Option-pricing theory may then be used to compute the value of the equity issued by the firm. The value of the corporate debt is then equal to the firm value less the value of the equity. An advantage of the model is that its components have natural interpretations, and the observable variables in the model can be used to impute default probabilities. A disadvantage of the model is that it is difficult to get it to apply to general coupon bonds and to capture realistic default triggers.

12.4 REDUCED-FORM MODELS

Reduced-form models are one of the first widely developed classes of corporate bond models. This class of model is ideally suited for actuaries and is an important part of the life insurance actuarial knowledge base, given the importance of corporate bonds to life insurers. These models are generally based on point process dynamics (Poisson process type mechanisms that actuaries are familiar with), which may be specified through a stochastic default intensity process in analogy to the force of mortality in life insurance mathematics. Reduced-form models may not allow for rating migration. The fundamental modeling objects are the time of default and the intensity process. (These have natural analogues to time of death and force of mortality.)

The simplest reduced-form class of models is based on the concept of an intensity process, and there is no explicit mechanism for rating change. This is the class of model that is developed in Duffie and Singleton (1999). While one may attempt to impute ratings for bonds modeled in this framework, it is the case that the intensity processes will have fixed parameters, which necessarily involve some degree of mean reversion. Consequently, the long-term credit behavior of an issuer name cannot deviate significantly from the parameters that the bond was issued with.

This creates significant practical issues in attempting to use this class of models when ratings migration is required. This class of models may be unsuitable for applications requiring realistic rating migration. Nonetheless, this class of models is very useful in many circumstances for which rating migration is not a first-order issue. For example, this class of models is useful for the modeling of certain classes of municipal bonds and sovereign bonds where rating changes are infrequent.

12.4.1 RISK-FREE TERM STRUCTURE AS A COMPONENT OF PRICING

Although we are interested in modeling credit risk, we also require the risk-free term structure of interest rates as a component of corporate bond pricing. Consequently, as we did in Chapter 11, we let r_t denote the usual instantaneous short rate of interest that is used in treasury bond pricing, and we allow for the short rate depending on state variables $x^{(1)}, x^{(2)}, x^{(3)}, \dots$

If one is using a three-factor treasury model, then one would have $r_t = x_t^{(1)} + x_t^{(2)} + x_t^{(3)}$.

We continue to let $v(t; T)$ denote the value at time t of one unit scheduled to be paid at time T , where $T > t$ and the scheduled payment is subject to credit risk. This notation is the credit risk analogue of $P(t; T)$ that was used in Chapter 11.

⁵ That is, either the bondholders will be paid the contractual redemption amount, or they will seize the assets of the firm and recover what they can.

12.4.2 CREDIT RISK AS AN INTENSITY PROCESS

Credit risk is measured and modeled by an intensity process for the default time. This process is denoted by h_t . One may think of the intensity process at time t as the instantaneous probability that the bond defaults by the next instant $t + dt$, conditional on the bond having survived to time t . In a simple deterministic model, the intensity plays the same role as to force of mortality in life insurance mathematics. This element of credit modeling is familiar to actuaries. Generally, the intensity is a nonnegative stochastic process with additional properties that are chosen according to the application at hand.

12.4.3 LOSS GIVEN DEFAULT

Also included in the model is the loss given default process, denoted by L_t . The loss given default process determines the consequence of a default at each point in time subject to the RMV protocol. Specifically, if default occurs at time t , then the bondholder receives payment in the amount of $(1 - L_t)v(t -, T)$ at time t , and the security is effectively wound up. For obvious reasons, $1 - L_t$ is referred to as the recovery rate. The notation $v(t -, T)$ serves to indicate the value of the bond at the instant prior to default.

12.4.4 ARBITRAGE-FREE PRICING WITH DIVIDEND

For a general intensity process and a nonzero recovery rate, it is not possible to immediately write down the price of the defaultable bond in terms of simple financial quantities (such as the price of an equivalent treasury bond and the survival probability of the credit risky bond). Instead, to derive the bond-pricing relation, one must leave the functional form of $v(t, T)$ unspecified and rely on the theory of arbitrage-free pricing with general dividend processes to derive a stochastic differential equation that the bond price must satisfy.

The end result of this analysis is the fundamental pricing relationship

$$(12.5) \quad v(t, T) = E_t^{\mathbb{Q}} \left[\exp \left(- \int_t^T (r_u + (1 - L_u)h_u) du \right) \right]$$

where the expectation must be taken under a risk-neutral measure and the subscript t denotes conditional expectation on the information available at time t . One may think of the risk-neutral measure as defining the risk premiums for credit risk.

Although Equation 12.5 is valid for a general stochastic loss given default process L_t , in practice one often assumes a constant value for this process. In some applications, one might use a Cox-Ingersoll-Ross square-root diffusion for the intensity process h . One might also use some of the more general intensity dynamics studied in Duffie and Singleton (1999).

Typically, one expects that Equation 12.5 will evaluate to a functional form such as

$$(12.6) \quad v(t, T) = F(t, T, \{x_t^{(i)}\}, h_t)$$

where $\{x_t^{(i)}\}$ serves to denote the factors that the treasury rate depends on.

Consequently, if the functional form is known either in closed form or from an efficient numerical procedure, then one may efficiently simulate bond prices by simulating the treasury state variables and the intensity process, as in Equation 12.6.

12.4.5 ZERO-RECOVERY BONDS

Corporate bond models with zero recovery are relatively simple to work with, because one may define the bond price processes directly as expectations of a known terminal random variable. For example, if r is the default time for the zero-recovery bond in question, then the bond price may be expressed as

$$v(t, T) = \mathbb{E}_t^{\mathbb{Q}} \left[\exp \left(- \int_t^T r_u du \right) \mathbf{1}_{\{\tau > T\}} \right].$$

This expression has the interpretation that the bond price is the discounted expected value of a unit at maturity when default does not occur.

The previous expression leads to an arbitrage-free model, since the discounted price process is a martingale under \mathbb{Q} :

$$\exp \left(- \int_0^t r_u du \right) v(t, T) = \mathbb{E}_t^{\mathbb{Q}} \left[\exp \left(- \int_0^T r_u du \right) \mathbf{1}_{\{\tau > T\}} \right].$$

When there is nonzero recovery under the RMV assumption, the gains process for the bond becomes more complicated, and there no longer exists a direct representation for the bond price in terms of a terminal random variable.

12.4.6 THE ROLE OF RISK PREMIUM

In simple reduced-form models, one can adjust the risk premiums to gain insight into how the model spread dynamics can be stressed. For CIR square-root diffusions, one can stress the model to an interesting level by adopting a highly skewed intensity process.

12.5 RATINGS-BASED MODELS

Rating-based models offer an intellectually appealing framework that appears to be the natural candidate for a workhorse corporate bond model. Indeed, rating-based models allow the user to prescribe a transition matrix and to model the changes in the spread of a corporate bond in response to changes in the underlying rating. While these are attractive features, these same features imply certain limitations on the robustness of the model output. For example, one no longer has idiosyncratic bond price behavior, since any two bonds with the same rating, maturity and coupon will evolve to have the same price.

We now discuss a general class of models that include those presented in Hurd and Kuznetsov (2007) and Feldhüttner and Lando (2008). A Heath-Jarrow-Morton (HJM) approach to the ratings-based framework is given in Eberlein and Ozkan (2003).

12.5.1 THE BASIC RATINGS-TRANSITIONS GENERATOR MATRIX

We allow for K rating classes $\{1, 2, \dots, K-1, K\}$, where state K is the default. As we have noted, we expect that in practice, the rating classes we will implement are {AAA, AA, A, BBB, HY, DEF}.

Fundamental to the analysis is the $K \times K$ -generator matrix for the rating transitions under the real-world measure \mathbb{P}

$$(12.5.1) \quad \mathcal{L} = \begin{pmatrix} \lambda_1 & \lambda_{1,2} & \lambda_{1,3} & \dots & \lambda_{1,K} \\ \lambda_{2,1} & \lambda_2 & \lambda_{2,3} & \dots & \lambda_{2,K} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda_{K-1,1} & \lambda_{K-1,2} & \lambda_{K-1,3} & \dots & \lambda_{K-1,K} \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

where $\lambda_{i,j} \geq 0 : i \neq j$ and $\lambda_i = -(\sum_{j \neq i} \lambda_{i,j}$, as is always the case for a true generator). The state of the Markov chain at time t is denoted by η_t .

The price at time t of a zero-coupon bond rated i at that time and scheduled to mature at time T is denoted $v(t, T)$. The model also includes factor processes that are used to drive the treasury term structure and to affect conditional changes in the generator matrix under the risk-neutral measure Q . This vector of factor processes is denoted by Z ,

and its components are denoted by $Z^{(i)}$. Stochastic calculus can be used to show that it is possible to induce a change of measure such that entries of the generator matrix are stochastic processes under the new measure.

In particular, one may find a change of measure that induces the generator of the form $L\mu(Z_t)$, where the new generator is multiplied by a scalar factor $\mu(Z_t)$ that depends on the factor processes. The precise form of $\mu(Z_t)$ need not be specified at this point in the model. Typically, it will be a linear function of the factor processes.

Let us write our basic generator matrix from (5.1) as

$$\mathcal{L} = \left(\begin{array}{c|c} L & v \\ \hline 0 \dots 0 & 0 \end{array} \right)$$

where $L \equiv [\lambda_{i,j}]_{i,j=1,\dots,K-1}$ is the $(K-1) \times (K-1)$ principal submatrix of \mathcal{L} and $v \stackrel{\text{def}}{=} (\lambda_{1,K}, \dots, \lambda_{K-1,K})'$. We may define a new generator as

$$(5.2) \quad \Lambda = L + \text{Diag}(\lambda_{1,K}, \lambda_{2,K}, \dots, \lambda_{K-1,K}) \equiv L + \text{Diag}(v).$$

The generator Λ defines a restriction of our original Markov chain to the non-default states $\{1, 2, \dots, K-1\}$. We will denote the states of this Markov chain by X_t .

12.5.2 RMV PRICING

Using the form for the intensity of a Markov chain, one may adapt the RMV pricing argument to show that

$$(5.3) \quad v^i(t, T) = E_t^{\mathbb{Q}} \left[\exp \left(- \int_t^T [r(Z_u) + L(u, Z_u, X_u)v(X_u)\mu(Z_u)] du \right) \middle| X_t = i \right].$$

Evidently, $v(X_u)\mu(Z_u)$ is the intensity process for the default event, and this intensity fluctuates as the Markov chain X changes states. The loss-given-default variable can depend on time, the state of the Markov chain X and the factor processes. We now consider the simplest version of Equation (5.3), which defines our model bond prices at time zero, namely

$$(5.4) \quad v^i(0, T) = E^{\mathbb{Q}} \left[\exp \left(- \int_0^T [r(Z_u) + L(u, Z_u, X_u)v(X_u)\mu(Z_u)] du \right) \middle| X_0 = i \right].$$

We may condition on \mathcal{F}_T^Z , the entire history of the factor process and consider the quantity

$$(5.5) \quad V(0, X_0) \stackrel{\text{def}}{=} E^{\mathbb{Q}} \left[\exp \left(- \int_0^T L(u, Z_u, X_u)v(X_u)\mu(Z_u) du \right) \middle| X_0, \mathcal{F}_T^Z \right]$$

together with the more general quantity

$$(5.6) \quad V(t, X_t) \stackrel{\text{def}}{=} E^{\mathbb{Q}} \left[\exp \left(- \int_t^T L(u, Z_u, X_u)v(X_u)\mu(Z_u) du \right) \middle| X_t, \mathcal{F}_T^Z \right].$$

Note that the only randomness remaining in Equation (5.5) is that of the Markov chain itself and that, conditional on the factor process Z , this is a time-inhomogeneous Markov chain.

Generalizing the calculation performed in Di Graziano and Rogers (2009, 60), we may show that the vector function $V(t; \bullet)$ satisfies the first-order differential equation

$$(5.7) \quad \frac{\partial V}{\partial t} = -\tilde{Q}_t^Z V$$

and boundary condition

$$V(T, \cdot) = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

where

$$(5.8) \quad \tilde{Q}_t^Z \stackrel{\text{def}}{=} [\Lambda - \text{Diag}(L(t, Z_t, 1)v(1), \dots, L(t, Z_t, K-1)v(K-1))] \mu(Z_t).$$

We may also use the notation

$$\tilde{\Lambda}_t^Z \stackrel{\text{def}}{=} [\Lambda - \text{Diag}(L(t, Z_t, 1)v(1), \dots, L(t, Z_t, K-1)v(K-1))].$$

Therefore, we may write equation (5.4) as

$$(5.9) \quad v^i(0, T) = \mathbb{E}^{\mathbb{Q}} \left[\exp \left(- \int_0^T r(Z_u) du \right) V(0, X_0) \right].$$

12.5.3 LOSS GIVEN DEFAULT

At this level of generality, Equation (5.9) is of little practical value, because we do not have available tractable expressions for $V(0, X_0)$. However, if we suppose that the loss-given-default random variables do not depend on the factor processes, then

$$(5.10) \quad \frac{\partial V}{\partial t} = -\tilde{Q}_t V$$

and boundary condition

$$V(T, \cdot) = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

where

$$(5.11) \quad \tilde{Q}_t \stackrel{\text{def}}{=} [\Lambda - \text{Diag}(L(t, 1)v(1), \dots, L(t, K-1)v(K-1))] \mu(Z_t).$$

The next level of specialization, and the one that leads to a closed-formula solution for our zero-coupon bond prices in Equation (5.3), is to suppose that recovery rates depend only on X . This then leads to

$$(5.12) \quad \frac{\partial V}{\partial t} = -\tilde{Q}V$$

and boundary condition

$$V(T, \cdot) = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

where

$$(5.13) \quad \tilde{Q} \stackrel{\text{def}}{=} [\Lambda - \text{Diag}(L(1)v(1), \dots, L(K-1)v(K-1))] \mu(Z_t).$$

It is convenient to use the notation

$$\tilde{\Lambda} \stackrel{\text{def}}{=} [\Lambda - \text{Diag}(L(1)v(1), \dots, L(K-1)v(K-1))]$$

so that we may write

$$\tilde{Q} = \tilde{\Lambda} \mu(Z_t).$$

Since

$$\tilde{\Lambda} \mu(Z_t) \int_0^t \tilde{\Lambda} \mu(Z_s) ds = \left(\int_0^t \tilde{\Lambda} \mu(Z_s) ds \right) \tilde{\Lambda} \mu(Z_t),$$

we have an explicit solution for $V(0, \cdot)$ as

$$(5.14) \quad \begin{bmatrix} V(0, 1) \\ V(0, 2) \\ \vdots \\ V(0, K-1) \end{bmatrix} = \exp \left(\int_0^T \tilde{\Lambda} \mu(Z_s) ds \right) \cdot \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

Therefore,

$$(5.15) \quad v^i(0, T) = \sum_{k=1}^{K-1} \mathbb{E}^{\mathbb{Q}} \left[e^{-\int_0^T r_s ds} \left[\exp \left(\tilde{\Lambda} \int_0^T \mu(Z_s) ds \right) \right]_{i,k} \right]$$

and more generally,

$$(5.16) \quad v^i(t, T) = \sum_{k=1}^{K-1} \mathbb{E}_t^{\mathbb{Q}} \left[e^{-\int_t^T r_s ds} \left[\exp \left(\tilde{\Lambda} \int_t^T \mu(Z_s) ds \right) \right]_{i,k} \right].$$

In order to make effective use of (5.16) we must diagonalize $\tilde{\Lambda}$ as

$$(5.17) \quad \tilde{\Lambda} = BDB^{-1}$$

where B is the matrix of eigenvectors and D is the usual diagonal matrix populated with the corresponding eigenvalues of $\tilde{\Lambda}$.

Since such a diagonalization is always possible if $\tilde{\Lambda}$ has distinct eigenvalues, we can always make a minor adjustment to $\tilde{\Lambda}$ if needed so as to ensure that Equation (5.17) holds. Ideally, the diagonalization occurs over real numbers, but one cannot guarantee that this happens. It is possible to handle the case for which $\tilde{\Lambda}$ is diagonalizable over complex numbers C but not over *real numbers* (i.e., some eigenvectors have nonzero imaginary parts), but one must code the evaluation of Equation (5.16) to recognize the special conjugate structure that occurs in this case.

At this point, the key to evaluating Equation (5.16) is to realize that

$$(5.18) \quad \exp \left(\tilde{\Lambda} \int_t^T \mu_s ds \right) = B \text{Diag} \left(d_1 \int_t^T \mu(Z_s) ds, \dots, d_{K-1} \int_t^T \mu(Z_s) ds \right) B^{-1}$$

which follows from the assumption in Equation (5.17) and the series expansion for the matrix exponential.

The last set of assumptions that are required to obtain a closed formula is to ensure that structure of $\mu(Z_s)$ permits the factor processes to be evaluated, including any interaction terms with the factors that are driving the treasury model. A typical example is to assume a linear structure for the function μ . The bond price formula may be partially simplified without explicit assumptions on μ as follows.

Let us use the notation $B = [b_{i,j}]$ and $B^{-1} = [b_{i,j}^{-1}]$

Then

$$\left[\exp\left(\tilde{\mathcal{L}} \int_t^T \mu_s ds\right) \right]_{i,j} = e'_i B \text{Diag}\left(d_1 \int_t^T \mu_s ds, \dots, d_{K-1} \int_t^T \mu_s ds\right) B^{-1} e_j$$

which simplifies to

$$\sum_{n=1}^{K-1} b_{i,n} b_{n,j}^{-1} \exp\left(d_n \int_t^T \mu_s ds\right)$$

Using this expression in (5.16), we find

$$v^i(t, T) = \sum_{k=1}^{K-1} \sum_{n=1}^{K-1} b_{i,n} b_{n,k}^{-1} E_{t,t,x}^Q \left[\exp\left(- \int_t^T r_s ds + d_n \int_t^T \mu_s ds\right) \right]$$

Under the assumption that r_s and μ_s are linear functions of the factor processes, one may be able to evaluate this expression in closed form using the moment generating functions associated with the factor processes.

Since both r_s and μ_s are affine functions of X and since the components of X are independent, this can be evaluated as a sum of products of CIR moment-generating functions.

12.6 OTHER ISSUES

Among other metrics, one should check that the model in use correctly captures spread distributions, total return characteristics, and the correct correlation with other asset classes. The validation of a corporate bond model or any other financial model is a critical part of the process of using such models in business applications.

For more on considerations relating to model validation, see Chapter 8.

12.6.1 WHAT TO DO ABOUT 2008?

The causes and effect of the events of 2008 remain unsettled. Just as uncertainty exists about the drivers of the Great Depression, so too will be the case with respect to the Great Recession. One is faced with the need to capture some features of systemic risk without destroying historical relationships within the credit risk model. Some advances are being made: for example, improvements to rating-based models to incorporate better spread dynamics and easier fitting of initial yield curves, improvements to estimation algorithms for rating-based models, and enhancements to existing corporate bond models that allow the features of the financial crisis to emerge in simulation but otherwise preserve the historical long-run model behavior.

12.7 CHAPTER REFERENCES

Cairns, A. 2004. *Interest Rate Models: An Introduction*. Princeton, NJ: Princeton University Press.

Di Graziano, G., and L. Rogers. 2009. A Dynamic Approach to the Modelling of Correlation Credit Derivatives Using Markov Chains. *International Journal of Theoretical and Applied Finance* 12(1): 45–62.

- Duffie, D. 2011. *Measuring Corporate Default Risk*. New York: Oxford University Press.
- Duffie, J. D., J. Pan and K. Singleton. 2000. Transform Analysis and Asset Pricing for Affine Jump-Diffusions. *Econometrica* 68(6): 1343–1376.
- Duffie, J. D., and K. Singleton. 1999. Modeling Term Structures of Defaultable Bonds. *Review of Financial Studies* 12(4): 687–720.
- Eberlein & Ozkan: The Defaultable Levy Term Structure - Ratings and Restructuring Mathematical Finance, Vol. 13, No. 2 (April 2003), 277-300
- Feldhütter, P., and D. Lando. 2008. Decomposing Swap Spreads. *Journal of Financial Economics* 88: 375–405.
- Huge, B., and D. Lando. 1999. Swap Pricing with Two-Sided Default Risk in a Rating-Based Model. *European Finance Review* 3: 239–268.
- Hurd, T., and A. Kuznetsov. 2007. Affine Markov Chain Model of Multifirm Credit Migration. *Journal of Credit Risk* 3(1): 329.
- Jarrow, R., D. Lando and S. Turnbull. 1997. A Markov Model for the Term Structure of Credit Spreads. *Review of Financial Studies* 10(2): 481–523.
- Jarrow, R., H. Li, S. Liu and C. Wu. 2010. Reduced-Form Valuation of Callable Corporate Bonds: Theory and Evidence. *Journal of Financial Economics* 95: 227–248.
- Jeanblanc, M., M. Yor and M. Chesney. 2009. *Mathematical Methods for Financial Markets*. New York: Springer.
- Lando, D. 1998. On Cox Processes and Credit Risky Securities. *Review of Derivatives Research* 2: 99-120.
- Schönbucher, P. 2003. *Credit Derivative Pricing Models: Models, Pricing and Implementation*. New York: John Wiley & Sons.
- Yu, F. 2007. Correlated Defaults in Intensity-Based Models. *Mathematical Finance* 17(2): 155–173.

Chapter 13: Equity Index Models

The preceding two chapters have looked at models for interest rates. The fundamental object that was modeled in those chapters was a bond, and from the price of the bond, an associated interest rate was derived. In this chapter, we will look at how to model the price of an equity index, together with its dividend payments, so that a model for total return can be obtained.

Equity index models are fundamentally different from bond models for a few reasons. These differences depend on which bond market we are comparing with equities. For the purposes of the high-level discussion here, we are thinking about a liquid equity market versus treasury or high-quality investment-grade bonds. The value of a bond is known on its maturity date, but the price of an equity index is random in perpetuity. Equity indices tend to be more volatile than most bond prices. The volatility of equity returns tends to show periods of elevated volatility and lower volatility, with the periods of elevated volatility clustering together while bond returns are more stable. Equity prices can exhibit sudden large changes over short periods of time that can be interpreted as jumps in the equity index price, but this is not typical for bonds.¹

Section 13.1 provides some general considerations of equity models. A realistic equity model should incorporate stochastic volatility, jump behavior and accurate return behavior. Equity models provide total return characteristics, often split between an ex-dividend pricing model component and a dividend yield model component. More robust models may interfere with the concept of market completeness, limiting the ability to conform to requirements of arbitrage-free pricing.

Section 13.2 provides a description of the standard Black-Scholes model, which is the traditional workhorse model for equity returns and associated derivatives pricing. A description of the mathematics underlying the model and its components is provided. The Black-Scholes model has the disadvantage of carrying a constant volatility term and a constant dividend yield.

Section 13.3 provides a description of the Heston model, which generalizes the Black- Scholes model by introducing stochastic volatility to produce more realistic equity return behavior and permit more accurate pricing of derivatives. The stochastic process incorporates two Brownian motion terms that allow for the correlation of the price and volatility.

Section 13.4 discusses the stochastic volatility with jumps (SVJ) model, which is an additional generalization of the Heston model by introducing a jump component to the price level. This can produce very accurate equity return distributions and path-wise price behavior, but it also requires complex mathematics to apply it effectively.

Section 13.5 provides a discussion and basic mathematics of the regime-switching model. This also builds off the Black-Scholes model but has the feature of changing parameter sets in a random fashion, comparable to ongoing shocks to the economy. This model is capable of significantly fattening the tails of stock price returns but can produce a market that is incomplete. Examples and mathematical notation are provided for stock price dynamics and option pricing. Discussion of pros and cons of the model also is provided.

Section 13.6 provides a brief discussion of econometric models. These models are particularly challenging in risk-neutral applications.

13.1 GENERAL CONSIDERATION OF EQUITY MODELS

The state of econometric and financial modeling knowledge has reached the point where quite realistic equity models are available. One can select from a range of models featuring jump behavior, stochastic volatility and feedback effects between volatility and jumps. It is possible to generate very robust path-wise behavior and accurate return distributions. However, the introduction of some of these advanced features does come with a cost: The asset market becomes incomplete. Incomplete means that all sources of risk cannot be hedged away by trading in the universe of available assets, and as a result, one cannot arrive at a unique price for some contingent cash flows by arbitrage-free pricing considerations alone. Therefore, one of the most important negative consequences

¹ It can be argued that interest rates (equivalently bond prices) do jump on occasion.

of using some of these advanced modeling techniques is that the asset market becomes incomplete and additional assumptions are needed to price derivatives. How this issue is resolved is beyond the scope of this chapter. If one is interested only in real-world risk simulation, then the incompleteness is not a concern. The challenges posed by incomplete models tend to be ignored, but one should keep in mind the two issues that incompleteness introduces:

- Some cash flows that are contingent on the equity index cannot be hedged by the assets available for trading.
- Some cash flows that are contingent on the equity index do not have a unique price based on the absence of arbitrage alone.

The latter condition can be overcome to some extent by adopting other methods for assigning a price. The former condition is more problematic if one is interested in hedging effectiveness, as would be the case in the risk management of a variable-annuity book.

13.1.1 WHAT MUST THE MODEL DO?

An equity index model must produce total returns for the equity index it is intended to model. Total returns consist of the combined effect of price returns and dividends. In some applications, it is required that the ESG separately produce the price return and dividend components of total return.

It is common for an ESG to need to model more than one equity index per economy. For example, an ESG for the United States may require equity indices for large-cap, mid-cap and small-cap returns. The simultaneous modeling of multiple equity indices requires that appropriate co-movements between the indices be allowed for. Co-movements between indices might be measured by correlation, and a standard way to correlate equity indices is to correlate the innovation (or error terms) that drive price movements. If the equity indices also include jump dynamics, then the possibility of simultaneous jump movements across all indices also needs to be allowed for.

Insurance products such as variable annuities require the pricing and hedging of complex guarantees. In such applications, the equity index model must be capable of generating risk-neutral scenarios for the pricing of contingent cash flows. The computation of hedge ratios (i.e., option Greeks) also may be required.

Equity index models typically specify the ex-dividend price of the index as one model component and the dividend yield as the other. The dividend yield is taken to represent the payout rate of the equity index as a percentage of its current price level. Modeling dividend yield instead of a cash dollar dividend payout has advantages in imposing arbitrage-free restrictions on the model. It has the disadvantage that dividend yield models do not reflect actual market dividend payout behavior as well as a cash dividend model would. For the purposes of this chapter, we assume that dividends are accommodated by a model for dividend yield.

13.1.2 JUMPS

Conventional wisdom suggests that equity markets are subject to jumps in prices and that equity market volatility also may exhibit jumps. By their nature, jumps are difficult to measure, and the empirical finance literature is not settled on the matter. For example, Christensen and Podolskij (2014) assert that “jumps in financial asset prices are often erroneously identified and are, in fact, rare events accounting for a very small proportion of the total price variation.” They further conclude, “In both theory and practice, traditional measures of jump variation based on lower-frequency data tend to spuriously assign a burst of volatility to the jump component. As a result, the true price variation coming from jumps is overstated. Our estimates based on tick data suggest that the jump variation is an order of magnitude smaller than typical estimates found in the existing literature.”

Regardless of the empirical findings, experience with the performance of models with and without jumps leads one to conclude that stochastic volatility models alone require more variation than seems natural in order to capture the tails of equity returns. Some combination of stochastic volatility and jumps tends to place the equity returns in the right place relative to historical benchmark data.

There seems to be a need to include a jump component in equity returns in order to obtain realistic real-world price dynamics. For some risk-neutral applications, the jump component may be omitted because it complicates the pricing of derivatives.

Jondeau, Poon and Rockinger (2007) provide a range of additional equity models beyond what is discussed in this chapter.

13.1.3 ARE EQUITY RETURNS PREDICTABLE?

It is tempting and natural to try to explain equity returns with observable variables. For example, one might think that a model with conditional mean, estimated off of things like GDP, interest rates and unemployment, would be a good starting point in getting a sound equity model. Unfortunately, the inclusion of observable economic variables does not significantly improve the explanatory power of the equity model. For this reason, most ESGs do not attempt to relate equity returns directly to macroeconomic effects.

In Section 13.6, we will note an econometric approach that can be taken if one wants to build in such relationships. It must be noted that although such an approach would build in a relationship, the explanatory power of the model will remain weak.

13.2 THE STANDARD BLACK-SCHOLES MODEL

The Black-Scholes model was the traditional workhorse model for equity returns and associated derivatives pricing. The basic model assumes that the stock index price level has a lognormal distribution and a constant dividend yield.

The real-world ex-dividend price of the equity index is modeled by the process

$$(2.1) \quad \frac{dS_t}{S_t} = (r_t + \lambda_S \sigma_S) dt + \sigma_S dW_t.$$

The parameters λ_S and σ_S are constant. One may sometimes see the model expressed as

$$(2.2) \quad \frac{dS_t}{S_t} = \mu_t dt + \sigma_S dW_t$$

but the two formulations are equivalent once arbitrage-free pricing restrictions are considered.

Comparing Equations (2.1) and (2.2), we arrive at the nice interpretation that

$$(2.3) \quad \mu_t = r_t + \lambda_S \sigma_S \quad \Rightarrow \quad \frac{\mu_t - r_t}{\sigma_S} = \lambda_S$$

where λ_S measures the risk-return trade-off in the equity index and is referred to as the market price of equity price return risk.

The dividend yield on the price index can be taken as a constant ∞ or can be modeled by a more general process such as

$$(2.4) \quad dD_t = \kappa_D (\theta_D + A_t - D_t) dt + \sigma_D \sqrt{D_t} dZ_t.$$

where κ_D , θ_D and σ_D are constant parameters. The process A_t may be included to cause dividend yield to fluctuate around other financial variables, examples being interest rates and GDP growth.

Correlation can be allowed for between movements in the equity price index and the dividend yield.

13.2.1 ON THE IMPLEMENTATION OF THE DIVIDEND YIELD

The dividend payments for the model are calculated based on the dividend yield rate as follows.

If one holds ϕ_t units of the S&P 500 index over the interval $[t, t + dt]$, then one receives the cash flow

$$\phi_t S_t D_t dt$$

over this interval. Therefore, the dividend yield is the rate at which the dividend cash flow is paid per dollar of investment in the S&P 500.

As an example, suppose that an investor holds \$1,000 in the S&P 500 index at time t . Mathematically, the condition that an investor holds \$1,000 in the index means that the number of units held in the index (denoted ϕ_t) times the unit price of the index (which we have denoted S_t) is equal to \$1,000, in other words:

$$1000 = \phi_t S_t.$$

If the current dividend yield is $D_t = 0.035$, then at time t , the investor is being instantaneously paid an amount equal to $1,000(0.035)dt = 35dt$.

13.3 THE HESTON MODEL

The Heston model (Heston 1993) generalizes the Black-Scholes model by introducing stochastic volatility. Instead of a constant σ_s parameter for price volatility, a stochastic process is used for volatility. This has the advantage of generating more realistic equity return behavior and permits more accurate pricing of derivatives.

Let S_t denote the equity index price level. The Heston model can be described by the following two general equations:

$$(3.1) \quad \frac{dS_t}{S_t} = \mu_t dt + \sqrt{V_t} dW_t^{(1)}$$

$$(3.2) \quad dV_t = \alpha(\theta - V_t)dt + \gamma \sqrt{V_t} dW_t^{(2)}$$

where the two Brownian motions $W^{(1)}$ and $W^{(2)}$ can be correlated with correlation parameter ρ , and the drift term μ_t may take on a fairly general stochastic form. The stochastic process V is the volatility process for the equity index. The process V fluctuates but remains nonnegative due to properties associated with its specified dynamics. Since V can trivially be set to a constant, the model nests Black-Scholes.

A nonzero value for the correlation parameter ρ will associate changes in volatility with price returns. For example, if $\rho < 0$, then an increase in volatility will tend to be associated with negative price return.

The dynamics can also be expressed in terms of the logarithm of the equity index price level, and this is commonly found in the literature. When the logarithm of the equity index price level is used, there is a small change in the dynamics of the model, but no practical differences result.

A very large literature has developed around this model that addresses the accurate simulation of returns and the precise computation of derivatives.

13.4 THE STOCHASTIC VOLATILITY WITH JUMPS (SVJ) MODEL

The SVJ model generalizes the Heston model by introducing a jump component to the price level. The model is also referred to as the Bates model (Bates 1996).

Let S_t denote the equity index price level. The SVJ model can be described by these two equations:

$$(4.1) \quad \frac{dS_t}{S_t} = \mu_t dt + \sqrt{V_t} dW_t^{(1)} + J_t$$

$$(4.2) \quad dV_t = \alpha(\theta - V_t)dt + \gamma \sqrt{V_t} dW_t^{(2)}$$

where J_t is a jump process and where, as with the Heston model, the two Brownian motions $W^{(1)}$ and $W^{(2)}$ can be correlated with correlation parameter ρ , and the drift term μ_t may take on a fairly general stochastic form.

The jump process can be taken to be of the compound Poisson type that is familiar to actuaries. In practice, the jump process usually incorporate a feedback effect, whereby the volatility process affects the likelihood of a jump occurring. This type of feedback effect is known as the leverage effect and is important in capturing the tendency for volatile equity returns to be accompanied by more large drawdowns.

The model has the advantage that it can produce very accurate equity return distributions and path-wise price behavior. It has the disadvantage that the pricing formulas are quite involved and require a high level of mathematical knowledge to effectively apply.

13.5 REGIME-SWITCHING MODEL

The regime-switching model is based on the standard Black-Scholes lognormal dynamics with the additional feature that these dynamics can change parameter sets in a random fashion. Such an underlying change is interpreted as ongoing unobservable shocks to the economy. For simplicity, we will restrict our discussion to a two-regime model. We are following the model as developed in Hardy's (2001) well-known paper.

13.5.1 STOCK PRICE DYNAMICS

There is a two-state Markov chain with monthly transition probabilities $p_{1,1}$, $p_{1,2}$, $p_{2,1}$, and $p_{2,2}$. We are labeling the two possible states as state 1 and state 2.

Evidently, two parameters are required to define this transition matrix, since $p_{1,1} + p_{1,2} = 1$ and $p_{2,1} + p_{2,2} = 1$. The state of the Markov chain at month k will be denoted by ρ_k .

The stock price process, denoted S_t , follows the real-world dynamics:

$$\frac{dS_t}{S_t} = \begin{cases} \hat{\mu}_1 dt + \sigma_1 dW_t^{(1)} & \rho_{12:t} = 1 \\ \hat{\mu}_2 dt + \sigma_2 dW_t^{(1)} & \rho_{12:t} = 2 \end{cases}$$

It is more convenient to write the model out as a month-on-month return (assuming one month is the finest simulation interval) formulation.²

The return for the n th month is (assuming time is expressed in years):

$$\log\left(\frac{S_{n/12}}{S_{(n-1)/12}}\right) = \begin{cases} \mu_1/12 + \sigma_1 \sqrt{1/12} \cdot \varepsilon_n^{(1)} & \rho_{n-1} = 1 \\ \mu_2/12 + \sigma_2 \sqrt{1/12} \cdot \varepsilon_n^{(1)} & \rho_{n-1} = 2 \end{cases}$$

where $\varepsilon^{(1)}$ are standard normal random variables (which can be correlated with their interest-rate counterparts, if desired). It is possible to allow for correlation between the driving Brownian motion $W^{(1)}$ and other Brownian motions in the ESG, such as those used to drive changes in interest rates. It is not possible to generalize the drift term (i.e., instantaneous expected return) if we wish to employ the type of estimation procedure given in Hardy (2001).

The simulation of the stock price process is straightforward. It requires the simulation of the Markov chain together with normal random draws weighted by the appropriate parameter set.

The form of the dividend yield process can be made as we have discussed in section 13.2.1. The presence of stochastic dividends complicates the closed formula for option pricing that one can obtain in this model, and to retain a tractable structure, we need some approximations to dividends in the option formula.

13.5.2 OPTION PRICING

Let us consider the basic option-pricing formula in the case where interest rates are constant (at the value ∞) and the stock does not pay dividends. The price of a European call option expiring at time T with strike price K is given by

² To make a strict translation between the two formulations, we should have $\hat{\mu}_j = \mu_j - \frac{1}{2}\sigma_j^2$.

$$\mathbb{E}^{\mathbb{Q}} \left[e^{-\delta T} (S_T - K)_+ \right].$$

Furthermore,

$$S_T = S_0 \cdot \exp \left((1/12) \cdot \sum_{k=1}^{12 \cdot T} \mu_{\rho_{k-1}} + \sqrt{1/12} \cdot \sum_{k=1}^{12 \cdot T} \sigma_{\rho_{k-1}} \cdot \varepsilon_k^{(1)} \right)$$

If we let R denote the number of months the regime spends in state 1 (and $p(r)$ the probability function for R), then the distribution of $\log(S_T/S_0)$ is normal conditional on R . Specifically:

$$\log(S_T/S_0) \sim N \left(R\mu_1/12 + (12 \cdot T - R) \cdot \mu_2/12, R\sigma_1^2/12 + (12 \cdot T - R) \cdot \sigma_2^2/12 \right).$$

Therefore,

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}} \left[e^{-\delta T} (S_T - K)_+ \right] &= \mathbb{E}^{\mathbb{Q}} \left[\mathbb{E}^{\mathbb{Q}} \left[e^{-\delta T} (S_T - K)_+ \mid R \right] \right] \\ &= \mathbb{E}^{\mathbb{Q}} [BSP(R)] \\ &= \sum_{k=1}^{12 \cdot T} BSP(k) \cdot p(k) \end{aligned}$$

where $BSP(r)$ denotes the usual Black-Scholes option price based on

$$\log(S_T/S_0) \sim N \left(r\mu_1/12 + (12 \cdot T - r) \cdot \mu_2/12, r\sigma_1^2/12 + (12 \cdot T - r) \cdot \sigma_2^2/12 \right).$$

Thus, we are able to compute an option price under our two-regime model as a weighted sum of the usual Black-Scholes formulas. Of course, we need to compute $P(r)$ up front. In summary, we have the following formula for option prices:

$$\text{Call Option Price} = \sum_{k=1}^{12 \cdot T} BSP(k) \cdot p(k).$$

If we now introduce dividends as a constant dividend yield, this formula would have the immediate generalization to

$$\text{Call Option Price} = \sum_{k=1}^{12 \cdot T} \widehat{BSP}(k) \cdot p(k).$$

where $\widehat{BSP}(r)$ is the well-known formula for dividend-paying stocks with the same mean and variance as before.

13.5.3 DATA AND ESTIMATION METHODOLOGY

The regime-switching model can be estimated using a classical maximum-likelihood approach or by using a Bayesian likelihood approach. Hardy discusses the classical estimation (EM-type algorithm). This requires a recursive computation of the likelihood components together with a search algorithm. The Bayesian approach can be implemented using a Markov chain Monte Carlo algorithm.

13.5.4 PROS AND CONS OF THE MODEL

The model is capable of significantly fattening the tails of stock price returns. It can increase the downside risk with no upside improvement as well. The model offers a tractable range of option-pricing formulas. Since Black-Scholes is the basis for the formulas, there is a natural intuition on the model's behavior, and it is relatively simple to work with.

A drawback of this model is that it is incomplete. However, this is also true of stochastic volatility models. The source of the incompleteness in the regime-switching model is that there is no asset available to hedge the risk of a regime

switch. The source of the incompleteness in models with stochastic volatility is that there is no asset available to hedge the risk of stochastic changes in equity volatility. For such incomplete models, there is a range of viable option prices that are consistent with the absence of arbitrage. Each model needs an assumption about the price of regime switch risk or volatility risk. The method offered for the regime-switching model assumes this risk is unpriced. The assumption that volatility risk is unpriced has also been made in the literature. These assumptions could be altered, but an alteration would result in technical adjustments to the option-pricing formulas.

13.6 ECONOMETRIC MODELS

Very good econometric models are available for capturing equity returns and equity return volatility. Due to challenges in using these models in risk-neutral applications, the most straightforward applications are to real-world simulation. However, versions of these models can be applied in risk-neutral simulations, and there is a considerable literature that takes many powerful econometric models over to the risk-neutral world.

A popular class of econometric models for equity returns is the GARCH(1,1) formulation. This model is defined as

$$(6.1) \quad r_t = x_t' \beta + u_t$$

$$(6.2) \quad u_t = \sqrt{h_t} v_t$$

$$(6.3) \quad h_t = \zeta + \delta h_{t-1} + \alpha u_{t-1}^2$$

where the random variables v_t are independent and identically distributed (i.i.d.) with zero mean and unit variance.

In this formulation, the vector x_t consists of observable financial variables that are used to define the conditional mean of the process. One might include changes in interest rate levels, GDP growth, oil prices and so on. The critical feature of the model is that the volatility structure accommodates periods of high and low volatility, and high volatility tends to be persistent. The model has the advantage that it may be estimated directly using maximum-likelihood techniques. It has the disadvantage that it is a discrete time model and is therefore not as flexible for general ESG simulations.

More sophisticated GARCH formulations can be adopted. In many cases, researchers have found the GARCH(1,1) to be sufficient to capture the key properties in the data. Also, a range of modifications can be made to this class of models to include jump behavior.

13.7 CHAPTER REFERENCES

- Bates, D. 1996. Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options. *Review of Financial Studies* 9(1): 69–107.
- Christensen, Kim, Roel C. A. Oomen and Mark Podolskij. 2014. Fact or Friction: Jumps at Ultra High Frequency. *Journal of Financial Economics* 114: 576–599.
- Hardy, M. 2001. A Regime-Switching Model of Long-Term Stock Returns. *North American Actuarial Journal* 5(2): 41–53.
- Heston, S. 1993. A Closed-Form Solution for Options With Stochastic Volatility With Applications to Bond and Currency Options. *Review of Financial Studies* 6(2): 327–343.
- Jondeau, E., S.-H. Poon and M. Rockinger. 2007. *Financial Modeling Under Non-Gaussian Distributions*. New York: Springer.

Chapter 14: International Considerations

There are many reasons that international considerations can enter into the application of an ESG. Frequently, an insurance company must deal with lines of business in a multinational context. ESG practitioners often refer to an ESG that simultaneously simulates more than one economy as *multi-economy*. Multi-economy can be considered as a synonym for “international.” If an ESG is used to aggregate risks across multiple economies, then all of the relevant economies would need to be available, together with the foreign-exchange rates necessary to combine the results in a single currency. Today’s investing environment is international in nature, as managers look to different regions of the world for growth and seek some of the benefits of global diversification. Larger insurance companies typically have business interests that span more than one country.

There are many practical issues in constructing and calibrating a multi-economy ESG. There are issues of data availability. Less developed economies tend to have a smaller menu of assets available for trading. In a less developed economy, the assets that are available for trading may not represent a liquid market, thereby making it hard to interpret the data that is available. Then there is the question of how changes in variables within the individual economies are related across economies in the multi-economy context. For example, it may be that changes in long-term interest rates in one economy tend to be accompanied by changes in long-term interest rates in some but not all of the other economies being modeled.

Section 14.1 discusses the differences between single- economy and multi-economy ESG modeling. Broadly speaking, moving to a multi-economy ESG imposes cross-economy calibration and validation requirements that are not a consideration in a single-economy context. The need for foreign-exchange rates is another distinction between single-economy and multi-economy ESGs. A multi-economy framework differs from single-economy modeling in the need for global correlation matrices and foreign-exchange processes that serve to link the constituent economies together into a global model.

Section 14.2 discusses differences between important variables across economies. Since the objective of an ESG is to capture market behavior, the financial variables that an ESG simulates for a given economy depend on the actual market structures associated with that economy. ESG variables that appear in several economies may have materially different payment or risk characteristics despite sharing similar names.

Section 14.3 addresses data issues in modeling global economies and associated challenges in setting calibration targets. One of the primary challenges in building multi-economy ESGs is that the availability of historical data varies greatly across economies. Some economies have long time series of data across a wide range of asset classes. Other economies may have short data histories for a small number of asset classes. Economies that are in transition may have available data series that need to be carefully interpreted in light of structural changes occurring in those economies.

Section 14.4 discusses the importance of capturing co-movements between financial variables across economies, and the limitations of what can be inferred from history. The simultaneous modeling of multiple economies requires that the components of each economy react sensibly as a global entity. There is no universally accepted way to measure co-movements of financial variables, but statistical correlation is one standard approach. Some of this can be achieved through correlation parameters, but some if this is structural and should be a part of the design of the dynamics of the global model. Two aspects of this general concept are the need to have a realistic proxy for global risk, and a distribution of global scenarios that is suitable for understanding the diversification effects of various business strategies.

Section 14.5 details the nature of exchange rate movements and their importance in multi-economy models. Foreign-exchange rates are the variables that complete the linkage between the component economies in a multi-economy ESG. Exchange rates are needed to understand the historical risk-return relationships between global asset classes, and exchange rates are needed to link the simulation results so that insurance risks and investment outcomes can be aggregated across economies.

Section 14.6 addresses the construction of a multi-economy ESG from existing single-economy models.

14.1 SINGLE-ECONOMY VERSUS MULTI-ECONOMY MODELING

Whether one is using a single-economy ESG or a multi-economy ESG, there are common considerations in putting together a solid ESG that we have touched on in earlier chapters. At the top of the list is the requirement that the ESG should produce realistic behavior for the simulated financial variables. In a single-economy context, “realistic behavior” encompasses things such as return distributions, stochastic volatility and co-movements between variables within the economy, such as equity returns and corporate bond returns. In a multi-economy context, “realistic behavior” encompasses these things as well but also requires that the correlation between interest rates across economies behave correctly and that the calibration targets across economies be coherent. Broadly speaking, moving to a multi-economy ESG imposes cross-economy calibration and validation requirements that are not a consideration in a single-economy context.

It is possible that a multi-economy ESG application might only look at business results within the constituent economies. For example, an insurer with business interests in the United States, Canada and the United Kingdom might want to look only at risk and profitability for the portions of their business pertaining to the United States, Canada and the United Kingdom. In such an application, realistic co-movements between financial variables across all three countries are of concern, in order for the model to correctly measure joint outcomes where results are unfavorable across all jurisdictions. It would not normally be sensible to simulate the three economies independently of one another, since that is not how these three economies typically relate to one another. Therefore, the need to capture co-movements between key financial variables across economies is an important additional consideration in the multi-economy ESG setting.

More comprehensive applications of multi-economy ESGs will involve aggregating positions in a common currency. Therefore, a multi-economy ESG will normally produce foreign-exchange rates as part of its output. Once exchange rates are available, the ESG becomes a truly global model in which asset prices directly relate to one another and international investing becomes available. The need for foreign-exchange rates is another distinction between single-economy and multi-economy ESGs. A detailed discussion of exchange rates is provided in Section 14.5.

If the multi-economy models are required to be arbitrage-free, then there are technical conditions that the traded assets across all component economies must simultaneously satisfy. These conditions are related to the arbitrage-free requirements for single-economy models, but they are more delicate to express, impose and verify, because the dynamics of all traded assets must be analyzed under a common currency. Thus, in a multi-economy framework, the arbitrage-free restrictions involve exchange rate processes, and the determination of risk-neutral dynamics is correspondingly more involved.

An important practical question is whether a multi-economy model needs to be constructed by simultaneously specifying all economies in the model, or whether it is possible to bootstrap up to a multi-economy ESG from existing single-economy models. If the answer is a qualified yes, then one can develop a list of single-economy models with their single-economy calibrations and select which of these are to be used to define a multi-economy simulation run, thereby giving the ESG structure a great deal of flexibility. In the yes case, if the parameters of one economy are changed, it would not affect the other economies, and changes in calibrations could be made to one economy that did not require a set of global parameter adjustments. If the answer is no, then one would have to construct a large global economy that includes all economies that might be required, and calibration would have to be performed on this global entity. Under a no, if an economy were needed that was not on the original list, then the global economy would have to be reconstructed and recalibrated.

Fortunately, the answer to the question is a qualified yes. The answer is qualified because one cannot obtain a global economic model by simply pasting together individual economies, but a coherent global economic model can be obtained by linking existing models for individual economies together using global correlation matrices and foreign-exchange processes. A detailed discussion of this is provided in section 14.6.

The perspective we take is that of an ESG practitioner working with continuous-time arbitrage-free models. If one departs from this classical framework, then other approaches for obtaining a global economic model from individual economies are possible. Thus, a multi-economy framework differs from single-economy modeling in the need for global correlation matrices and foreign-exchange processes that serve to link the constituent economies together into a global model.

14.2 DETAILS OF SIMILAR ASSET CLASSES CAN VARY SIGNIFICANTLY ACROSS ECONOMIES

There is a core commonality to asset classes across global economies. Since the objective of an ESG is to capture market behavior, the financial variables that an ESG simulates for a given economy depend on the actual market structures associated with that economy. For example, each economy is expected to include a market for treasury bonds, equities and a broad inflation measure. Many economies will have corporate bond markets. Some economies will have mortgage-backed securities (MBSs). Other mortgage-related assets, like the German Pfandbriefe, can be important to a given economy. Real estate investments may be made through different vehicles, such as real estate investment trusts (REITs), as opposed to brick-and-mortar properties. An ESG may model the dominant investment vehicle for a given asset class, rather than include multiple variations of the same asset class. There also can be unique asset classes that are not present in other economies.

Institutional details can vary significantly across economies. European sovereign bonds pay annual coupons, while U.S. Treasury bonds pay semiannual coupons. U.S. corporations pay quarterly dividends on equity shares, which are equal across quarters unless the dividend is changed, while European equity shares tend to pay annual dividends with a possible interim payment. Large liquid markets for MBSs exist in the United States and to a lesser extent in Canada and Denmark. However, the prepayment features that are critical to the behavior of U.S. MBS pass-throughs are not present in the Canadian mortgage market, and the available interest rate lock-in period for Canadian mortgages is much shorter than the 30 years that is available in the U.S. market. Therefore, the characteristics of MBS securities in the U.S. and Canadian markets are very different. Pfandbrief securities, jumbo Pfandbriefe in particular, represent an important liquid asset class for Germany but are of much less importance elsewhere in the world.¹ Some government bond markets do not have securities with long maturities, Brazil being one such example. Other government bond markets issue securities with very long maturities; two examples are Great Britain and France.

In summary, there can be significant differences between the universe of ESG variables across economies. Some economies will have a lot of asset classes available while other economies will have a much slimmer menu to choose from. ESG variables that appear in several economies may have materially different payment or risk characteristics despite sharing similar names.

Table 14.1
ILLUSTRATIVE ASSET CLASSES AND KEY ESG VARIABLES FOR THREE DEVELOPED ECONOMIES

United States	Germany	United Kingdom
Treasury bonds	Treasury bonds	Treasury bonds
Equity indices	Equity indices	Equity indices
Corporate bonds	Corporate bonds	Corporate bonds
Inflation-indexed bonds	Inflation-indexed bonds	Inflation-indexed bonds
Mortgage-backed securities		
Municipal bonds	Jumbo Pfandbriefe	
Inflation	Inflation	Inflation
GDP	GDP	GDP
Unemployment	Unemployment	Unemployment
REITs	Properties	Real estate

¹ Outside of Germany, the corresponding asset class is best described as covered bonds.

14.3 DATA ISSUES IN MODELING GLOBAL ECONOMIES

Whether one is building a single-economy ESG or a multi-economy ESG, data is essential for informing the development, calibration and validation processes that are central to an ESG. One of the primary challenges in building multi-economy ESGs is the availability of data. The availability of historical data varies greatly across economies. Some economies, such as the United States and the United Kingdom, have long time series of data across a wide range of asset classes. Other economies may have short data histories for a small number of asset classes. Economies that are in transition may have available data series that need to be carefully interpreted in light of structural changes occurring in those economies.

A key implication of all of this is that even when the relevant data for each economy has been identified, the aggregated global data is ragged. Ragged data means that we may have one series for a variable that runs from January 1950 to current (i.e., through today's date), but the best we can get for another series is data from April 1994 to current. Ragged also means that we may have one series for a variable that runs from January 1967 to current, but the best we can get for another series is data from January 1967 to June 2012. This latter case can happen when asset classes cease to exist, as happened to data for AAA and AA corporate bonds for many economies in recent years.²

Ragged data presents challenges in setting calibration targets. For example, when data is ragged, one cannot simply average returns across a common historical window to obtain coherent average return targets across economies.³ As another example, when recent data is missing but the ESG needs to continue to model that variable, then the practitioner is forced to make an extrapolation of the previous data against the relevant remaining global benchmarks.

Economies with developing financial markets can have relatively large asset classes that are important to the companies domiciled there, but for which there is very limited or no historical data. The Chinese financial bond market, bank deposit market and private debt markets are examples of this situation. The relative risk-return characteristics for such asset classes may be hard to reconcile, particularly within an arbitrage-free modeling framework. For example, in the Chinese market, financial bonds might be thought of as default-free, with a better average yield than for other assets also viewed as default-free. Obviously, one cannot have an arbitrage-free model in which some default-free assets are dominated by other default-free assets. This situation can develop for many reasons, one of which might be that financial bonds are driven by government policy and not market demand. Nonetheless, this type of situation shows up in global data, and one must then deal with the resulting model implications. A global model may attempt to include these asset classes, but model development for such asset classes cannot proceed as it would for a well-developed economy. One may also be faced with doubts about the quality or accuracy of official statistics. A classic example is concerns about the accuracy of official GDP statistics for China, as discussed in Holz (2014).

Several very useful references are available that deal with international capital markets. Ibbotson and Brinson (1993) survey global capital markets and provide an overview of the returns of major asset classes as they stood in 1993. Dimson, Marsh and Staunton (2002) provide another valuable in-depth look at global capital markets and their risk-return characteristics. While this reference is now more than 10 years old, it is supplemented each year by the Global Investment Returns Yearbook report put out by these and other authors through Credit Suisse.

14.3.1 INTERPRETING DATA FOR ECONOMIES IN TRANSITION

One of the primary uses for historical data is the setting of ESG calibration targets. As we have noted, the process of setting calibration targets for a multi-economy ESG is often complicated by the fact that the availability of historical

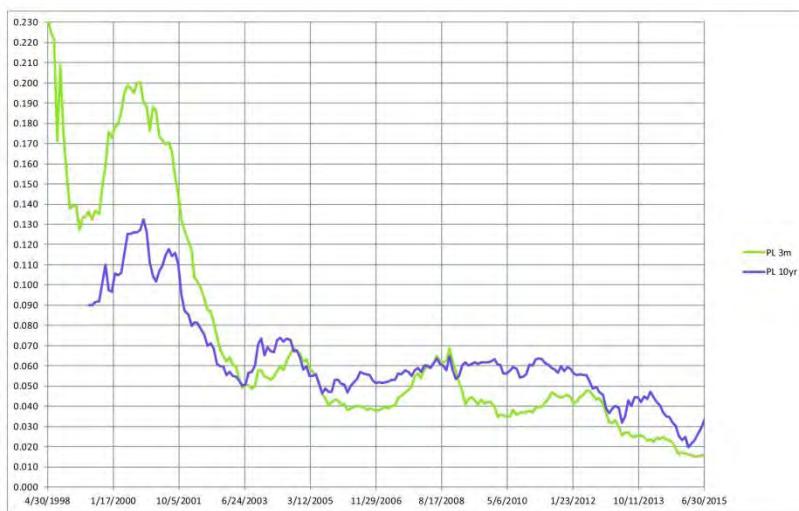
² Post financial crisis, AAA and AA corporate bond series were discontinued, since there were too few bonds remaining to reliably assess representative interest rate curves.

³ Even if data was available for the same historical periods across all economies, there are reasons why average returns across a common historical window might not be the right choice for average return targets.

data varies greatly across economies. Related to this is the fact that, for emerging economies, only a portion of the available data may be relevant to current modeling. Let us consider the case of setting calibration targets for Poland.⁴

Poland is one example of an emerging economy where an econometric analysis of the available data is insufficient to construct a sound set of calibration targets. Indeed, as Figure 14.1 shows, the data on Polish interest rates begin when the country was recovering from its Communist epoch, which officially ended in 1989, and Polish interest rates have moved significantly ever since. The recent era of central bank manipulation calls into question the interest rate data over the past few years, not just for Poland but globally. One thing that is clear from Figure 14.1 is that Polish interest rate dynamics have undergone a significant transition.

Figure 14.1
POLISH THREE-MONTH AND 10-YEAR TREASURY RATES



Sources: Bloomberg; Conning, Inc.

It may be reasoned that Poland's geographical location, robust manufacturing sector and high education level make it probable that Poland will eventually converge to similar economic performance as Germany's. Therefore, it might be appropriate that the calibration targets for Poland are comparable to the targets used for Germany. There is some expert judgment involved in the rate at which this convergence is assumed to happen, but this is not part of the calibration targets. Alternatively, one might expect that Poland is likely to have a comparable but somewhat lesser financial standing to Germany's, and that long-term calibration targets for Poland should therefore be at a small positive spread to Germany. The convergence between German and Polish interest rates has been appreciable in recent years. The key issues to note here is that calibration targets for Poland must be established relative to other global economies, and some expert judgment is unavoidable.

It should be noted that the flavor of this example is not unique to emerging economies. If an ESG practitioner takes U.S. Treasury interest rate data from 1970 to present and attempts to set calibration targets for U.S. interest rates, it will immediately be apparent that there has been some kind of change in U.S. interest rate behavior that commenced in the early 1990s. This practitioner's task will be further confounded by the peculiar behavior of interest rates post financial crisis, for which there are now seven years of actual data to consider. This practitioner is likely to appreciate

⁴ Calibration targets are approached differently, depending on the nature of the economy that we are setting the targets for. There are three broad cases:

Case 1: Data is available for the economy for the entire calibration window, and it is believed all available data is relevant. Statistics for data over the entire calibration window could be used for setting targets.

Case 2: Data is available for the economy for a portion of the calibration window, and it is believed all available data is relevant.

Econometric analysis is needed in relating the available data to the full calibration window. Some expert judgment may be needed.

Case 3: It is believed that some of the available data is not relevant. Econometric analysis and expert judgment are needed.

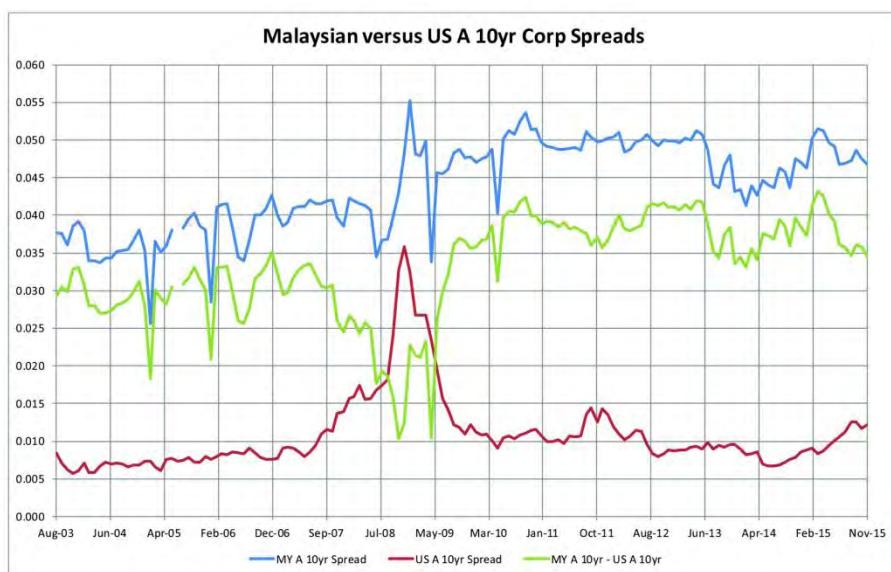
that the high interest rate data from the 1970s and 1980s are less relevant to the setting of current calibration targets.

14.3.2 SIGNIFICANT DIFFERENCES IN MARKET NORMS

Another interesting example of the challenges in multi-economy modeling is presented by the Malaysian corporate bond market. As we know from our discussion of corporate bonds, ratings are designed to reflect expected credit costs associated with holding a particular security. In an international context, one might expect that the credit costs associated with holding an A-rated bond in one economy would be similar to the credit costs associated with holding an A-rated bond in a different economy. Often, this is true. Sometimes it is not, as we will now illustrate for Malaysia.

Based on Bloomberg data, the average spread to treasury on an A-rated 10-year bond for the period August 2003 through November 2015 for U.S. industrial corporate bonds is 108 basis points and for Malaysian corporate bonds it is 439 basis points. In the U.S. market, a rated bond with a 439-basis-point spread to treasury would almost certainly have a below-investment-grade rating. Figure 14.2 illustrates the situation.

Figure 14.2
COMPARISON OF SPREADS ON U.S. AND MALAYSIAN CORPORATE BONDS



Sources: Bloomberg; Conning, Inc.

One might feel that a consequence of the large Malaysian spread differential is that Malaysian corporate bonds must have significant credit costs if their risk-return characteristics are consistent with other A-rated bonds in a global context. Another explanation for this wide spread differential might involve liquidity effects or differences in what the ratings mean. In any event, this feature of the Malaysian corporate bond market has significant modeling implications for the ESG practitioner.

A key observation is that similar-sounding asset classes can have very different behavior across economies.

14.4 THE IMPORTANCE OF CO-MOVEMENTS BETWEEN ECONOMIES

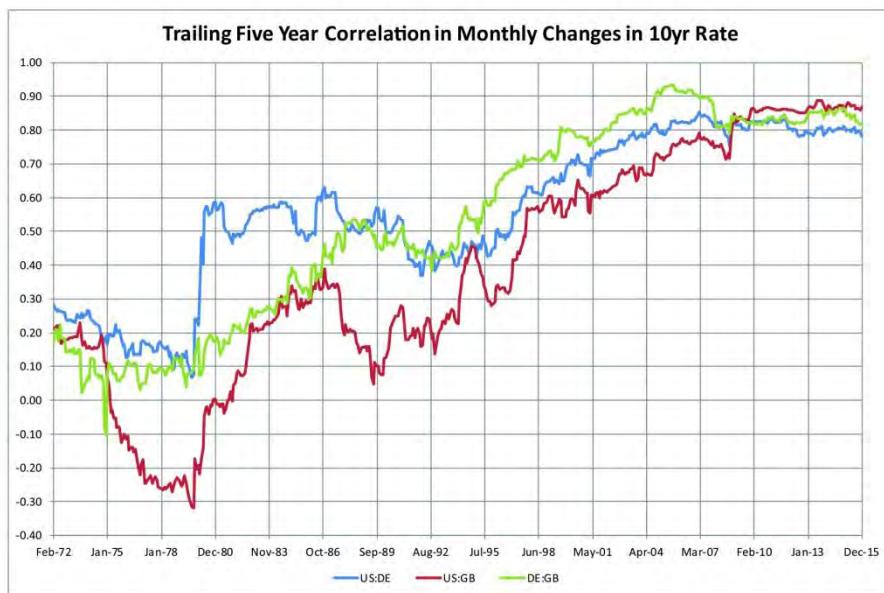
The simultaneous modeling of multiple economies requires that the components of each economy react sensibly as a global entity. In a multi-economy ESG context, it is natural to ask whether the co-movements of variables across economies are reasonable and consistent with historical norms. As always, the devil is in the details of what constitutes reasonable and what the historical norms really are. One viewpoint would be that the ESG output should encompass history and produce extreme but plausible outcomes. Explicit calibration targets might also serve as a reasonableness check. It is far more difficult to draw definite conclusions from the historical data concerning the

relationships between the co-movements of variables that can be used to establish historical norms for a multi-economy framework.

There is no universally accepted way to measure co-movements of financial variables, but statistical correlation is one standard approach.⁵ Figure 14.3 shows the five-year trailing correlations of the monthly changes in the 10-year treasury rate for the United States, Germany and the United Kingdom, using data from March 1967 through December 2015.

Figure 14.3

CROSS-ECONOMY CORRELATIONS FOR MONTHLY CHANGES IN 10-YEAR TREASURY RATES



Sources: Bloomberg; and Conning, Inc.

Figure 14.3 shows that there has been a significant change in the way interest rates for the United States, Germany and the United Kingdom co-move, with the trend being toward a convergence of these markets. Furthermore, the correlations in recent years have been relatively stable. Figure 14.3 might be interpreted as showing that correlations move around significantly over time and that there are periods of time for which correlations can be relatively stable. The implications for ESG modeling are subject to a range of interpretations. Some practitioners would conclude that using static (i.e., constant) correlation parameters in a global ESG model is adequate. Others might decide that a dynamic correlation structure is warranted.

More sophisticated dynamic correlation structures are a challenge to fit into a tractable arbitrage-free pricing framework. Furthermore, it is very difficult to identify the drivers of the changes in correlations and to map these drivers into fundamental ESG variables that can be modeled. In practice, ESGs that try to incorporate dynamically changing correlations are likely to rely on latent variables (i.e., sources of randomness with no direct economic interpretations) to drive changes in correlation. Consequently, some practitioners might prefer the easier-to-explain assumption of constant correlation.

Buhler and Zimmermann (1996) find that, in the case of German and Swiss interest rates, neither volatilities nor correlation coefficients are stable (constant) over time, but that the correlations are much more stable than the volatilities. They note that volatilities may be unstable even for rather long subperiods. This type of finding seems to be typical of interest rate behavior across economies.

⁵ We are referring to the usual concept of correlation, also known as Pearson's correlation coefficient. Other measures of correlation/association are possible but are not considered here.

The simultaneous modeling of multiple economies requires that the components of each economy react sensibly as a global entity. Some of this can be achieved through correlation parameters, but some if this is structural and should be a part of the design of the dynamics of the global model.

Correlation is probably the most direct tool that can be used to measure and model co-movements of ESG variables. A common technique for accommodating correlation in an ESG structure is to correlate the innovation terms (i.e., error terms) that drive changes in the relevant financial variables. Specific correlation dynamics in the innovation terms can be induced by using a Choleski factorization of the relevant correlation matrix. Mathematically, if the correlated standard error terms are to be the vector Z with correlation matrix P , then these can be generated as LW , where $P = LL^T$ is the Choleski factorization of P . Other techniques can also be used to capture correlation structures, copulas being a popular alternative.

It is very difficult to explain why correlations are unstable and to identify the cause(s) of changes in the way variables relate to one another across economies. In Dimson, Marsh and Staunton (2002, 114), the general observation is made that “correlations depend on the underlying structural relationship between countries, and these have naturally changed over time. Major shifts have occurred in world power blocks and politics through decolonization, wars being replaced by peace, economic growth and development, shifting trade flows, economic unions, changing currency regimes, and so on.”

Intuitively, a cause of sudden changes in movements could be flight to quality—the idea that when market risks become elevated, the best course of action is to sell all but the highest-quality assets and move into strong currencies. Cho et al. (2016) discuss aspects of this effect and conclude that it creates positive correlation between currency and equity in emerging markets and negative correlation between currency and equity in developed markets. These effects are difficult to capture in an ESG and are all the more difficult to model if one wishes to retain an arbitrage-free model structure. If one is willing to depart from an arbitrage-free framework, then a wide range of econometric tools can be applied, one such approach being dynamic conditional correlation models, as in Engle (2009).

A deeper study of the finance literature does not lead to firmer conclusions about the nature and causes of co-movements of financial variables across economies that could be used to inform the global dynamics for an ESG. Longin and Solnik (1995) study the correlation of returns across major economies over the period 1960–1990. They find that the correlation matrices are unstable over time. Longin and Solnik (2001) find that correlation increases in bear markets but not in bull markets. Solnik, Boucrelle and Le Fur (1996) note that International correlations for stocks and bonds fluctuate widely over time, and that international correlation increases in periods of high market volatility. A general discussion of correlation and its importance in financial models may be found in Meissner (2015). A high-level discussion of how the breakdown of correlation and its modeling created large losses in credit products is given in Triana (2009, Chapter 4).

Despite the complexities of the way correlations behave, it should be remembered that ESGs are fairly general models about broad pieces of the economy. Indeed, ESGs model things like large-cap equity returns, investment-grade bond returns, inflation and GDP. At the other extreme would be models of small-cap health care returns or returns on B-rated utility bonds for a manager interested in investment decisions at a very granular level. At such a granular level, the intricacies of the correlations between returns would be a key consideration in developing an investment strategy. In the much more aggregated world of risk management, it is less important that ESGs capture the intricacies of global correlation structures than that they provide a platform that offers a plausible proxy for the manner in which global economies interact. Therefore, if an ESG can offer a reasonable correlation structure it is likely to be sufficient for most risk management applications.

Ensuring that a multi-economy ESG correctly captures the co-movement of variables across economies is important. Two aspects of this general concept are the need to have a realistic proxy for global risk and a distribution of global scenarios that are suitable for understanding the diversification effects of various business strategies.

14.4.1 CREATING A REALISTIC PROXY FOR GLOBAL RISKS

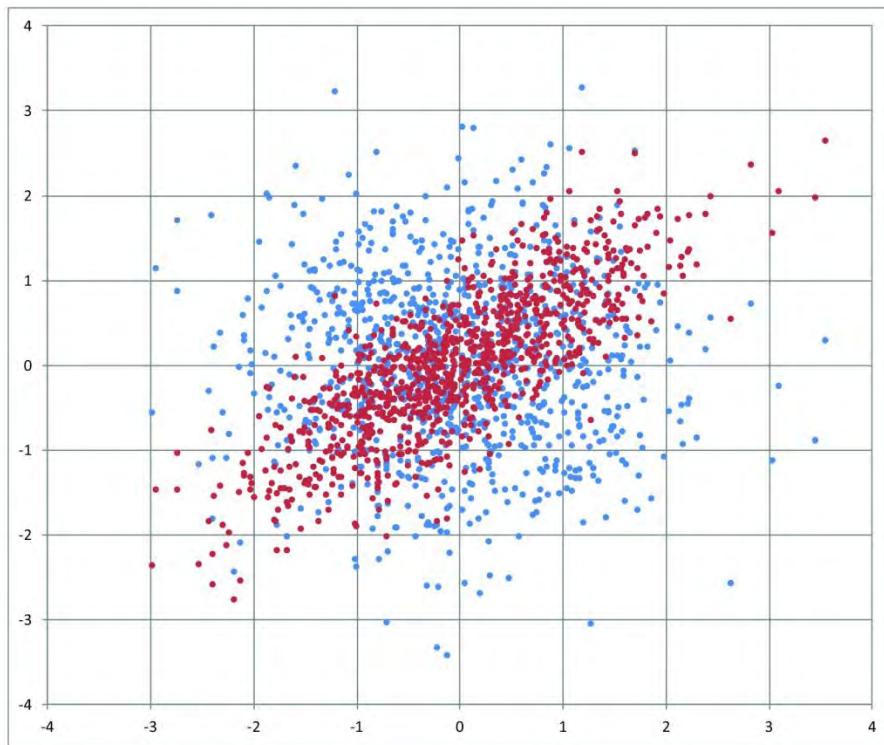
The identification of severe risk scenarios for an international insurer depends on the ESG being capable of simulating extreme outcomes occurring simultaneously across some or all of the economies in which the insurer operates. If

economies were simulated independently of one another, then the ESG output would lack scenarios where extreme conditions simultaneously occur in multiple economies. While bad outcomes could happen simultaneously across multiple economies even though the economies were simulated independently, the frequency with which this would occur would be greatly reduced from what would be observed in a correlated multi-economy setting.

Figure 14.4 illustrates the point. Suppose that the red pairs are a reasonable proxy for the risk profile we are concerned with. In this figure, the components of the blue pairs are independent, while the components of the red outcomes are correlated. The marginal distributions of all pairs are standard normal, so both the blue and red outcomes correctly capture the risk characteristics of the individual components of the pairs. The blue outcomes do not generate the same occurrences of jointly extreme outcomes that we see in the red outcomes. Additionally, the blue outcomes wind up generating a lot of simulated scenario pairs that reside in a part of the outcome space we do not expect to encounter (under our assumption that the red pairs are a reasonable proxy for the risk profile we are concerned with).

Figure 14.4

SIMULATED PAIRS OF STANDARD NORMAL RANDOM VARIABLES, UNCORRELATED AND CORRELATED.



Note: The blue pairs are simulated standard normal random variables (U_1, U_2), and the red pairs are constructed from these draws as $(U_1, \rho U_1 + \sqrt{1 - \rho^2} U_2)$. $\rho = 0.7$.

Source: Conning, Inc.

In practical terms, the lesson is that using independent economies as a proxy for a multi-economy framework will create a lot of irrelevant ESG scenarios. The flip side of the coin is that having too much correlation is also undesirable, as it will constrain the range of simulated outcomes relative to the outcomes that we consider plausible, and it will produce some extreme scenarios that we may not regard as plausible. Again using Figure 14.4 for illustration, if the blue pairs are a reasonable proxy for the risk profile we are concerned with, then using the red outcomes as a proxy for this risk profile would miss the important outcomes in the northwest and southeast corners of the chart and produce some extreme scenarios in the northeast and southwest corners of the chart that we do not expect to encounter.

In summary, an ESG that assumes too much independence may produce many scenarios of limited relevance to what the insurance company expects to encounter and fail to produce scenarios that reflect jointly occurring extremes across economies. Conversely, imposing strong views on correlation can limit the range of scenarios that are necessary for a fair sampling of the risk profile an insurance company faces and produce too many scenarios that reflect jointly occurring extremes in the global economy.

14.4.2 DIVERSIFICATION BENEFITS OF BUSINESS STRATEGIES

Insurers may want to use ESG scenarios to understand the diversification benefits of changes in lines of business across economies, the effects of a potential acquisition, changes in international reinsurance treaties, or the consequences of varying global allocations in their investment strategies. Getting the co-movement of variables across economies right is an essential part of measuring the benefits of diversification. In an investment context, it is the correlation structure that is important in understanding diversification effects. As noted in Ibbotson and Brinson (1993, 229), "An asset's diversification value depends on the correlation of its return with those of other assets." Correlation among global risk factors can limit the benefits of diversification. Ensuring that the co-movement of key financial variables across economies is realistic will provide insight into the limitations of risk mitigation strategies. Correctly capturing negative correlation effects can help insurers better position their risks with no cost to overall performance. Of course, if a multi-economy ESG is able to create a realistic proxy for global risks, then it follows that the application of these ESG scenarios will be suitable for understanding diversification effects. However, it is unlikely that any ESG will be given such a broad level of acceptance. In practice, one may be forced to accept some imperfections in the way that a multi-economy ESG handles the co-movement of variables across economies. Developing an understanding of the implications of various correlation assumptions and how they affect the outcomes associated with changes in business strategy is an important step in deciding which cross-economy features of an ESG can be accepted and which will require modification.

14.5 FOREIGN EXCHANGE (FX)

Foreign-exchange (FX) rates are the variables that complete the linkage between the component economies in a multi-economy ESG. Exchange rates are needed to understand the historical risk-return relationships between global asset classes, and exchange rates are needed to link the simulation results so that insurance risks and investment outcomes can be aggregated across economies.

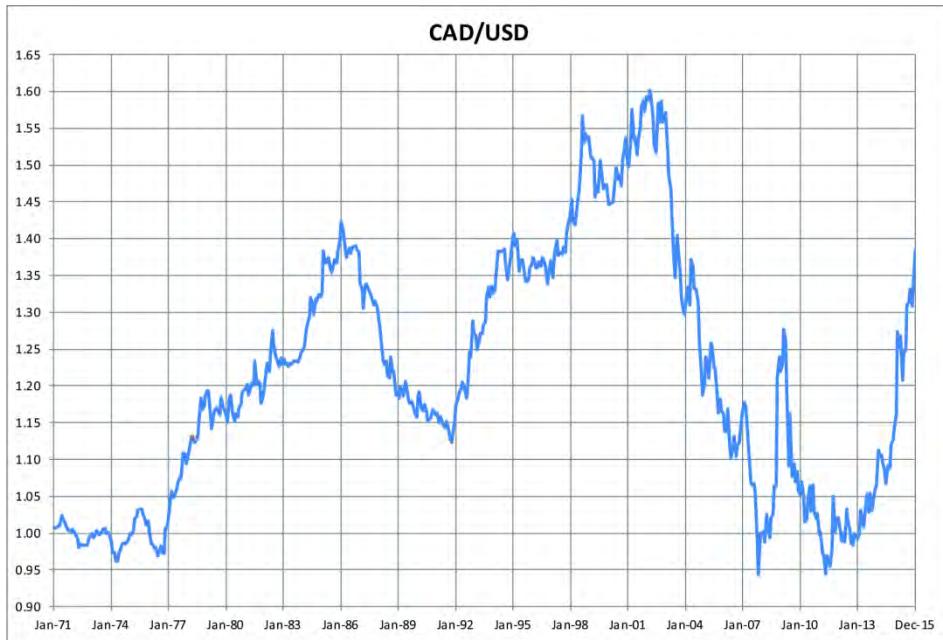
Useful overviews of currency markets and exchange rates may be found in Ibbotson and Brinson (1993, Chapter 2) and Dimson, Marsh and Staunton (2002, Chapter 7).

It is only in the relatively recent times, beginning from the early 1970s, that exchange rates have been floating. Therefore, we have data that goes back no more than about 40 years from which to understand FX rates. For economies that have experienced an economic trauma such as high inflation and have recently recovered, we have even less data. Nonetheless, global foreign-exchange markets are very liquid with high volumes of trade, and a lot can be inferred about the qualitative statistical features of exchange rates.

14.5.1 SOME OBSERVATIONS ON FX RATES

Two notable features of exchange rates are that they are subject to extended periods of trending and significant fluctuations in volatility. Whether there is mean reversion in exchange rates is open to debate, and the conclusions depend upon the statistical tests applied and the historical period examined. As Figure 14.5 shows, the relationship between the Canadian dollar (CAD) and the U.S. dollar (USD) usually seems to be in some type of trend. Furthermore, the CAD/USD exchange rate has historically traded in a band of about \$0.94 to \$1.60.

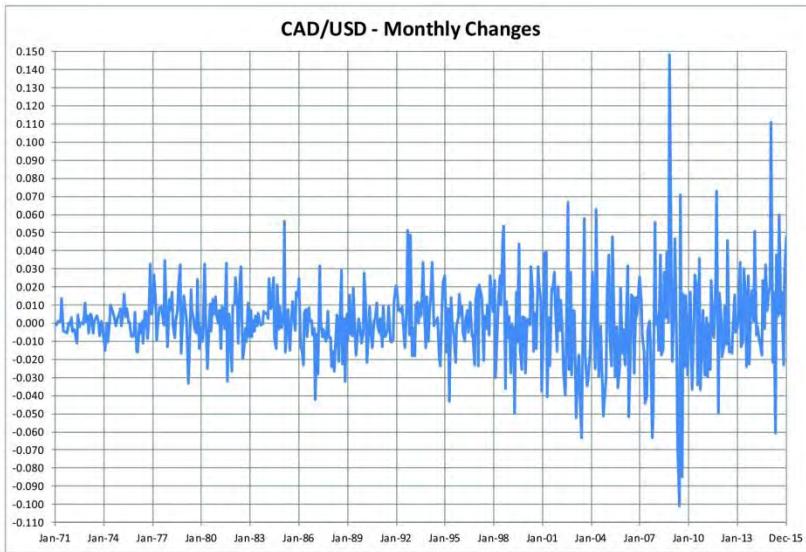
Figure 14.5
MONTHLY EXCHANGE RATES OF CANADIAN DOLLAR (CAD) TO U.S. DOLLAR (USD)



Sources: Bloomberg; Conning, Inc.

Figure 14.6 shows that the volatility of the CAD/USD exchange rate fluctuates over time.

Figure 14.6
MONTHLY CHANGES IN CAD/USD EXCHANGE RATE: EVIDENT HETROSCEDASTIC BEHAVIOR

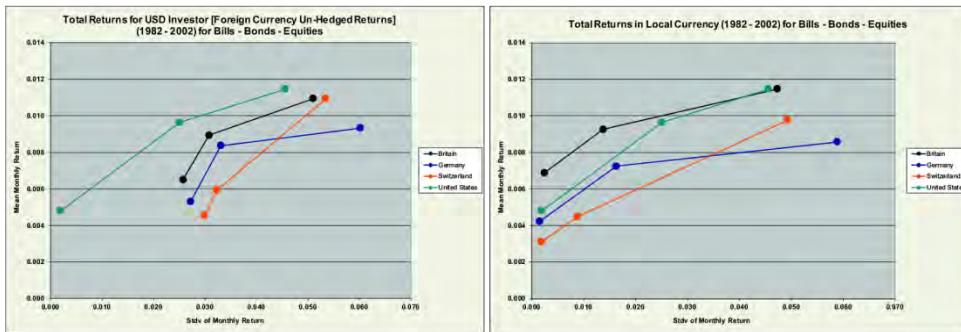


Sources: Bloomberg; Conning, Inc.

Exchange rates are essential in the calculations that investors need to perform in order to evaluate global investment opportunities. For example, Figure 14.7 shows the returns and standard deviations from treasury bills, treasury bonds and equities for the United States, Great Britain, Germany and Switzerland in local currency and U.S. dollars. Local-

currency returns are what an investor in that country would experience, while USD returns are what a U.S. investor would experience.

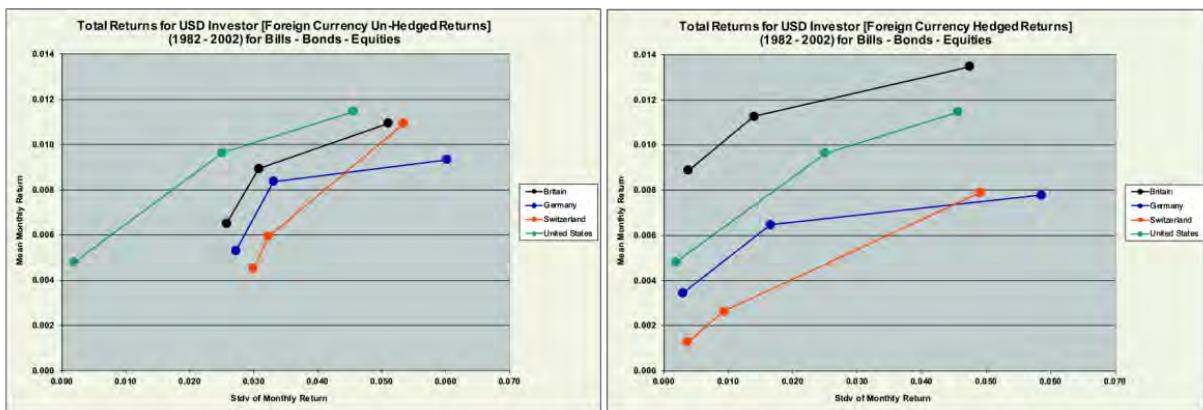
Figure 14.7
USD RETURNS VS. LOCAL-CURRENCY RETURNS



Sources: Bloomberg; Conning, Inc.

In practice, investors may seek to hedge some portion of their exchange rate risk. Figure 14.8 shows that this can produce quite different results from the unhedged case. Note how the risk-return structure changes when exchange rates are accounted for.

Figure 14.8
GLOBAL RETURNS FOR A USD INVESTOR, WITH CURRENCY RISK HEDGED AND UNHEDGED



Sources: Bloomberg; Conning, Inc.

Hedging of exchange rate risk can be done a variety of ways and may involve the use of derivatives contracts. If an ESG is to be used to evaluate a full range of global investment options including the hedging of exchange rate risk, then the ESG will need to include an arbitrage-free model for exchange rates.

14.5.2 MODELS FOR FOREIGN-EXCHANGE (FX) RATES

There are many sophisticated econometric models that can be used to capture the statistical features of exchange rate data. This literature is vast, and the reader may consult Johnson and Schneeweis (1994), Cuthbertson and Nitzsche (2004), Cheung and Erlandsson (2005) and Erdemlioglu, Laurent and Neely (2016), to name just a few representative resources. In some cases, a range of financial or macroeconomic variables may be included in the exchange rate model to attempt to explain the movements in the exchange rate, and a general econometric framework with explanatory variables is reviewed in Pilbeam and Langeland (2015). It is appealing to attempt to develop models that explain exchange rate movements. However, it can be argued that the incorporation of such

detailed mechanisms is not the way an ESG should work, since the focus of an ESG is the generation of plausible risk scenarios, not forecasting. The use of a robust econometric model may be suitable for some real-world ESG applications.

If one requires an arbitrage-free multi-economy model, there are additional technical considerations. To begin with, the arbitrage-free framework requires that one consider all traded securities simultaneously. As we discussed earlier, the nature of the absence of arbitrage is that there is no way to generate a free lunch from trading in available securities. In a multi-economy ESG, the securities available for trade consist of the universe of all traded securities from each constituent economy, compared in a single currency by application of the appropriate exchange rates.

For example, in a two-economy model for the United States and Canada, we could view the universe of all traded securities as the securities available for trade in the United States and the securities available for trade in Canada converted to U.S. dollars by the USD=CAD exchange rate. Let us denote the USD=CAD exchange rate at time t by D_t . Evidently, the CAD=USD exchange rate at time t is $1/D_t \equiv D_t^{-1}$. For each economy, United States ($i = 1$) or Canada ($i = 2$), let the money market account be denoted by $B_t^{(i)}$, the price of a zero-coupon bond be denoted by the usual notation $P^{(i)}(t, T)$. Many assets will be available for trade in each economy, and all must be considered simultaneously in the arbitrage-free framework. For illustrative purposes in the following example, we will assume that only money market accounts and zero-coupon bonds are available for trade.

With the exchange rate process defined, we can now bring assets from each economy over for trade in the other economy. Consequently, we can define a global asset market in the units of currency of either local economy. This is indicated in Table 14.2. Note that in the U.S. column, the assets that involve the symbol D_t are Canadian bonds denominated in U.S. dollars, and that in the Canada column, the assets that involve the symbol D_t^{-1} are U.S. bonds denominated in Canadian dollars.

Table 14.2
GLOBALLY TRADED ASSETS IN CURRENCY OF EACH ECONOMY

US (Economy 1)	Canada (Economy 2)
$B_t^{(1)}$	$D_t^{-1} B_t^{(2)}$
$D_t B_t^{(1)}$	$B_t^{(2)}$
$P^{(1)}(t, T)$	$D_t^{-1} P^{(1)}(t, T)$
$D_t P^{(2)}(t, T)$	$P^{(2)}(t, T)$

The no-arbitrage restrictions can be expressed in terms of either economy, and there is symmetry in the way that an equivalent martingale measure is selected. The key consideration is that the risk-neutral parameterization must simultaneously take into account all traded assets across all economies. The presence of the exchange rate process and the necessity of considerable notation make the framework more burdensome than what is needed to understand risk-neutral in a single economy. An approachable introduction to these ideas may be found in Baxter and Rennie (1996, Chapter 6).

A popular continuous-time model for exchange rates that can be readily applied in risk-neutral is a variation on the Heston model that we looked at for equity returns. If we denote the exchange rate by x_t , then

$$(5.1) \quad dx_t = \kappa(\mu - x_t)dt + \sqrt{V_t}dW_t^{(1)}$$

$$(5.2) \quad dV_t = \alpha(\theta - V_t)dt + \gamma\sqrt{V_t}dW_t^{(2)}$$

where the two Brownian motions $W^{(1)}$ and $W^{(2)}$ can be correlated with correlation parameter ρ . The parameter μ is sometimes replaced with the differential between the foreign and domestic interest rate, depending on the dynamics that are emphasized. The model is able to capture varying degrees of mean reversion and stochastic volatility.

14.6 CONSTRUCTING GLOBAL ECONOMIES FROM SINGLE-ECONOMY MODELS

The construction of an arbitrage-free multi-economy ESG requires that all models for traded assets fit together according to the restrictions that the arbitrage-free condition imposes. In theory, one could obtain a global arbitrage-free model by laying out all of the economies to be modeled and carefully obtaining a global risk-neutral measure. In principle, one can build in very intricate relationships between variables across economies and still have an arbitrage-free model, so long as one can impose the required conditions.

In practice, ESGs must add (or remove) economies over time, and the asset classes that are included within each economy may get expanded over time. It would be very inconvenient if one had to go back and reassess the global risk-neutral measure each time a modest change is made to the ESG. Furthermore, a given application may not call for all economies, and at times only some of the available economies will be simulated. For these reasons, global ESGs are not normally constructed based on the imposition of many detailed relationships between variables across economies. To be sure, provisions should be made to allow for appropriate correlations and some structural linkages across economies, but enough flexibility needs to be maintained so that portions of the ESG can be run alone and so that modest changes can be made to the ESG without requiring a total overhaul of its structure.

It turns out that it is possible to construct individual economies and then assemble these economies into a global arbitrage-free model, providing certain restrictions are imposed on the dynamics across economies. The fine details of what these restrictions are and why they permit a “modular” approach to a global ESG requires looking at the technical details of a global ESG and is beyond the scope of this chapter. However, the basic idea is straightforward, and we will give a high-level description of one method that can be used to retain a “modular” approach to the construction of a global ESG.

The phrase “driving Brownian motion” or “innovation term” is used to refer to the stochastic process that changes ESG variables. For instance, in Equation (5.1) in the prior section,, $W^{(1)}$ is the innovation term for the exchange rate level, and in Equation (5.2), $W^{(2)}$ is the innovation term for the volatility process of the exchange rate. Within each economy, there will be a collection of innovation terms associated with the models for that economy. The innovation terms within each economy will drive changes in the interest rates, equity returns, inflation, etc. for that economy. The foreign-exchange processes that link the economies will also have innovation terms, just as we have noted for Equations (5.1) and (5.2). If the innovation terms within each economy are not shared across other economies and if the innovation terms in the foreign-exchange processes are not shared across economies, then one can construct a global risk-neutral measure. The manner in which a global risk-neutral measure is constructed is based on bootstrapping up from the risk-neutral dynamics within each of the single economies, and it is this feature that gives the global ESG a degree of modularity. The condition that the innovation terms within each economy are not shared across other economies should not be taken to mean that we require that the innovation terms across economies be independent. On the contrary, the innovation terms will generally be correlated across economies. What the condition captures is the idea that the current state and future movements within each economy are determined by innovations that are modeled within that economy.

14.7 CHAPTER REFERENCES

- Baxter, M., and A. Rennie. 1996. *Financial Calculus: An Introduction to Derivative Pricing*. New York: Cambridge University Press.
- Bühler, A., and H. Zimmermann. 1996. A Statistical Analysis of the Term Structure of Interest Rates in Switzerland and Germany. *Journal of Fixed Income* 6 (December): 55–67.
- Cheung, Y., and W. Erlandsson. 2005. Exchange Rates and Markov Switching Dynamics. *Journal of Business and Economic Statistics* 23(3): 314–320.
- Cho, J., J. Choi, T. Kim and W. Kim. 2016. Flight-to-Quality and Correlation Between Currency and Stock Returns. *Journal of Banking and Finance* 62: 191–212.
- Cuthbertson, K., and D. Nitzsche. 2004. *Quantitative Financial Economics: Stocks, Bonds and Foreign Exchange*. 2nd ed. New York: John Wiley & Sons.

- Dimson, E., P. Marsh and M. Staunton. 2002. *Triumph of the Optimists: 101 Years of Global Investment Returns*. Princeton, NJ: Princeton University Press.
- Engle, R. 2009. *Anticipating Correlations: A New Paradigm for Risk Management*. Princeton, NJ: Princeton University Press.
- Erdemlioglu, D., S. Laurent and C. Neely. 2016. Which Continuous-Time Model Is Most Appropriate for Exchange Rates? *Journal of Banking and Finance* 61: S256–S268.
- Flesaker, B., and L. Hughston. 1996. Positive Interest: Foreign Exchange. In Hughston (1996).
- Holz, C. 2014. The Quality of China's GDP Statistics. *China Economic Review* 30: 309–338.
- Hughston, L., ed. 1996. *Vasicek and Beyond: Approaches to Building and Applying Interest Rate Models*. London: Risk Publications.
- Ibbotson, R., and G. Brinson. 1993. *Global Investing: The Professional's Guide to the World Capital Markets*. New York: McGraw-Hill.
- Ibbotson, R., and L. Siegel. 1991. The World Bond Market: Market Values, Yields, and Returns. *Journal of Fixed Income* 1(1): 90–99.
- Johnson, G., and T. Schneeweis. 1994. Jump-Diffusion Processes in the Foreign Exchange Markets and the Release of Macroeconomic News. *Computational Economics* 7: 309–329.
- Kritzman, Mark P. 2000. *Puzzles of Finance: Six Practical Problems and Their Remarkable Solutions*. New York: John Wiley & Sons.
- Longin, F., and B. Solnik. 1995. Is the Correlation in International Equity Returns Constant: 1960–1990? *Journal of International Money and Finance* 14(1): 3–26.
- . 2001. Extreme Correlation of International Equity Markets. *Journal of Finance* 56(2): 649–676.
- McCulloch, J. 1975. Operational Aspects of the Siegel Paradox. *Quarterly Journal of Economics* 89(1): 157–169.
- Meissner, G. 2015. It's All About Correlation. *Wilmott*, Issue 80 (November): 30–37.
- Melino, A., and S. Turnbull. 1990. Pricing Foreign Currency Options with Stochastic Volatility. *Journal of Financial Economics* 45: 239–265.
- Pilbeam, K., and K. Langeland. 2015. Forecasting Exchange Rate Volatility: GARCH Models Versus Implied Volatility Forecasts. *International Economics and Economic Policy* 12: 127–142.
- Rogers, L. C. G. 1997. The Potential Approach to the Term Structure of Interest Rates and Foreign Exchange Rates. *Mathematical Finance* 7: 157–176.
- Siegel, J. 1972. Risk, Interest Rates and the Forward Exchange. *Quarterly Journal of Economics*, 86(2): 303–310.
- . 1975. Reply: Risk, Interest Rates and the Forward Exchange. *Quarterly Journal of Economics*, 89(1): 173–175.
- Solnik, B., C. Boucrelle and Y. Le Fur. 1996. International Market Correlation and Volatility. *Financial Analysts Journal* 52(5): 17–34.
- Triana, P. 2009. *Lecturing Birds on Flying: Can Mathematical Theories Destroy the Financial Markets?* New York: John Wiley & Sons.

Annotated Bibliography

Here we list the references that are especially important for ESG modeling and applications. The citations within each section are in order of publication date. Some annotations cite papers that are not in this annotated collection. Reference data for those citations is at the end.

INTEREST RATES

Fair, R. C., and B. G. Malkiel. 1971. The Determination of Yield Differentials Between Debt Instruments of the Same Maturity. *Journal of Money, Credit and Banking* 3(4): 733–749.

The purpose is to study yields on three types of bonds: long-term U.S. government bonds, high-quality utility bonds and high-quality industrial bonds. This study strongly supports the theory that yield differentials of such bonds of the same maturity are influenced by the supply and anticipated flow of new issues. However, the authors also suggest that utility, industrial and government bonds are not perfect substitutes. The study is based on data for U.S. bonds traded in 1968–1969.

Cox, J. C., et al. 1985. A Theory of the Term Structure of Interest Rates. *Econometrica* 53: 385–407.

This paper has become a classic in the literature on the term structure of interest rates. The underlying bonds are default-free. For example, this model might be used to simulate returns on U.S. Treasury bonds. The underlying model is a general equilibrium model. The authors' basic model (now called the CIR model) has a single-state variable subject to three assumptions. The result is that the spot interest rate r satisfies

$$dr_t = \kappa(\theta - r_t)dt + \sigma\sqrt{r_t} dz_t$$

where z_t is a Wiener process and κ , θ and σ are positive constants. The model has "empirically interesting" properties:

- Negative interest rates are precluded.
- If the interest rate reaches zero, it can bounce back and become positive again.
- There is a steady-state distribution for the interest rate.

The authors explain many other features. The probability density at time s , conditional on its value at $t < s$, can be written explicitly in terms of a Bessel function. They provide formulas for the first two moments of the distribution of $r(s)$, conditional on $r(t)$ for $t < s$. They also derive formulas for zero-coupon bond prices and European call options on zero-coupon bonds.

The authors consider generalizations of the basic model, but the basic CIR model is the main contribution.

Ho, T. S. Y., and S.-B. Lee. 1986. Term Structure Movements and Pricing Interest Rate Contingent Claims. *Journal of Finance* 41(5): 1011–1029.

The authors describe an arbitrage-free interest rate model. The initial complete interest rate terms structure is given. Subsequence evolution through discrete time yields a complete arbitrage-free structure at each step. Applications include valuing options on (default-free) bonds and pricing callable bonds.

Pedersen, H. W., et al. 1989. Arbitrage-Free Pricing of Interest Rate Contingent Claims. *Transactions of the Society of Actuaries* 41: 231–279.

The authors discuss the pricing of bond options and interest-sensitive cash flows by discrete time binomial lattice models. The main contribution is an arbitrage-free binomial model that allows the initial term structure of interest rates to be prescribed exogenously. This allows the model price for each stream of fixed and certain cash flows to be set as the market price. Their approach can be viewed as a generalization of Ho and Lee (1986).

Tilley, J. A. 1990. A Stochastic Yield Curve for Asset/Liability Simulation. *Actuarial Approach for Financial Risks: First AFIR International Colloquium*, Paris, April 23–27, International Actuarial Association.

The author describes an empirical approach to long-term asset/liability simulations. The model has dynamic U.S. Treasury yield curves but no other variables. It is intended for modeling U.S. life insurer or pension assets and liabilities.

The yields are based on the bid prices for “on the run” U.S. Treasury securities, observed once each four weeks during December 16, 1981, to August 16, 1989, a total of 101 observed yield curves. The yield range (three months to 30 years) is mapped to the unit interval [0,1]. Orthogonal polynomials are fit to the 101 (transformed) yield curves on [0,1]. Each of the 101 yield curves was fitted to a third-order polynomial, resulting in a vector for $t = 0, 1, \dots, 100$. It turns out that the dynamics is described by an autoregressive process of order two, with a_t depending on a_{t-1} and a_{t-2} . Tilley reports two important results:

- The fitted coefficients exhibit mean reversion, with the result that the yield curves tend to revert over the long term to a normal, positively sloped form centered near 8.3%.
- The residuals have fat-tailed distributions that are explainable by a mixture of two normal distributions.

Heath, D., et al. 1992. Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation. *Econometrica* 60: 77–105.

This paper is important because it provided a unifying framework for understanding arbitrage pricing theory of default-free bonds and their contingent claims. The authors claim that all previous models are special cases, including Vasicek (1977), Brennan and Schwartz (1979), Ho and Lee (1986), and others.

The model takes the initial forward rate curve as given. The authors specify a general continuous-time stochastic process for its evolution across time, in such a way that it has an equivalent martingale probability measure. Under these conditions, markets are complete, and contingent claim valuation is then straightforward. This is an important theoretical paper, but it has no applications and only two simple examples.

Tilley, J. 1992. An Actuarial Layman’s Guide to Building Stochastic Interest Rate Generators. *Transactions of the Society of Actuaries* 44: 509–538.

This is a long discussion of important concepts in the theory of arbitrage-free pricing of interest rate contingent cash flows: arbitrage, complete markets, expected value pricing in the risk-neutral world, and stochastic interest rate models. The paper is written for actuaries, but it has no mathematical descriptions of the models. Indeed, there are exactly two equations in the entire paper.

A quote from the author’s concluding remarks:

This paper has shown that it is possible to gain a practical understanding of key concepts in financial economics without having to resort to a study of the mathematics of stochastic processes. Based on the assumptions that opportunities for riskless arbitrage do not exist and that the financial markets are complete, it was shown that a theory could be developed for pricing interest-rate-contingent cash-flow streams relative to the prices of all zero-coupon bonds, which are taken as exogenous inputs to the theory. The arbitrage-free prices are calculated in a straightforward manner by the expected-present-value algorithm, a technique that lies at the heart of actuarial science.

The comments at the end of the paper and the author’s responses are also valuable contributions. This paper won the 1992–93 Society of Actuaries Annual Prize and the Halmstad Prize for 1992.

Hull, J., and A. White. 1993. One-Factor Interest Rate Models and the Valuation of Interest-Rate Derivative Securities. *Journal of Financial and Quantitative Analysis* 28(2): 235–254.

This important paper discusses different approaches to arbitrage-free models of the term structure of interest rates. The authors present a numerical procedure for constructing one-factor models that are Markov and consistent with the initial term structure.

One approach is to specify a process for the short rate that has one or two unknown functions of the drift and then estimate the functions so that the process fits the initial term structure. The authors cite several examples of this approach, including models by Vasicek (1977), Cox et al. (1985) and Black and Karasinski (1991).

The authors also describe a general procedure for this procedure based on a trinomial tree. The tree can also be used to value options on bonds and interest rate options.

Goldstein, R. S., and W. P. Keirstead. 1997. On the Term Structure of Interest Rates in the Presence of Reflecting and Absorbing Boundaries. Available at SSRN: <http://ssrn.com/abstract=19840> or <http://dx.doi.org/10.2139/ssrn.19840>

The authors use the historic interest rate behavior in Japan in 1991–1995 to motivate introduction of a class of models that “admit the possibility of high volatility of low interest rates, while precluding the possibility of negative rates. This is accomplished by imposing reflecting or absorbing boundaries” on interest rates. They compare this to Black’s approach, which treats interest rates as options (1995).. Introducing barriers seems better because it allows closed-form solutions for bond prices and some derivatives. It also allows for maximum-likelihood estimation of parameters. The authors give a convincing argument for their approach versus Black’s . In Section 5 they illustrate the barrier approach when the spot rate process under risk-neutral dynamics is one of the following:

- (1) Brownian motion with drift
- (2) An Ornstein-Uhlenbeck process
- (3) An extended Cox-Ingersoll-Ross process
- (4) Longstaff dynamics: $r(t) = [x(t)]^2$, where $x(t)$ is Brownian motion with negative drift and a reflecting barrier at $x(t) = 0$
- (5) Multifactor where $r(t) = x_1(t) + \dots + x_n(t)$ is the sum of independent processes $x_k(t)$, each one of the types (1) to (4)

Delbaen, F., and H. Shirakawa. 2002. An Interest Rate Model With Upper and Lower Bounds. *Asia-Pacific Financial Markets* 9(3-4): 191–209.

The authors introduce a new interest rate model for the spot rate $r(t)$:

$$dr(t) = \alpha(r_\mu - r(t))dt + \beta\sqrt{[r(t) - r_m][r_M - r(t)]}dW(t)$$

where $W(t)$ is standard Brownian motion, $\alpha > 0$, $\beta > 0$ and $0 < r_m < r_M$ are constants. According to the authors, this type of process is called a Jacobi process, and it has been well studied in genetics. An important feature is that the spot rate is bounded away from zero, since it always satisfies $r_m < r(t) < r_M$. However, we do not usually see models with an upper bound on the spot rate. As the authors note, “The use of such models could be questionable.” The authors obtain some interesting results:

- (1) The CIR model can be obtained as a limiting case in which the lower bound is zero and the upper bound tends to infinity.
- (2) The authors show how the transition probabilities of the spot rate can be calculated as series expansions in Jacobi polynomials.
- (3) They show how to compute moments of the spot rate.
- (4) Assuming the spot rate dynamics is the pricing dynamics, they show how to compute prices of zero-coupon bonds.

There are no examples, and there is no attempt to estimate the model.

Gorovoi, V., and V. Linetsky. December 2003. Shadow Interest. *Risk*, pp. 81–84.

The authors study the Black (1995) approach, in which the spot rate r_t is the maximum of zero and the shadow rate x_t . They assume the shadow rate follows a Vasicek process. The main contribution is the detailed solution for zero-coupon bond prices based on an eigenfunction expansion. This is necessarily an approximation. They calibrate the model to Japanese bond market prices. Evidently, the same approximation can be applied to evaluate prices of options on zero-coupon bonds.

Takamizawa, H., and I. Shoji. 2003. Modeling the Term Structure of Interest Rates with General Short-Rate Models. *Finance and Stochastics* 7(3): 323–335.

Takamizawa, H., and I. Shoji. 2004. On the Accuracy of the Local Linear Approximation for the Term Structure of Interest Rates. *Quantitative Finance* 4(2): 151–157.

Miller, S., and E. Platen. 2005. A Two-Factor Model for Low Interest Rate Regimes. *Asia-Pacific Financial Markets* 11(1): 107–133.

The authors provide an interest rate term structure model for markets where one observes prolonged periods of low interest rates, such as Switzerland in the late 1970s and Japan in 1995–2005. The model has a savings account, B_t , and a growth optimal portfolio (GOP), S_t . The proposed model is an extension of the minimal market model derived previously by Platen. Here briefly are the dynamics:

$$dB_t = r_t B_t dt$$

$$dS_t = S_t \left[r_t dt + \sum_{k=1}^2 \theta_t^{(k)} (\theta_t^{(k)} + d\theta W_t^{(k)}) \right]$$

Initial values $B_0, S_0 > 0$ are given. The component $\theta_t^{(k)}$ is a nonrandom function of time, called the k th market price of risk with respect to the standard Brownian motion W^k ($k = 1, 2$). Note also that W^1 drives the short rate r_t , and W^2 is independent of the short rate.

The World Stock Accumulation Index (WSAI), denominated in Japanese Yen (JPY), is taken as a proxy for the GOP. The authors use observations of the WSAI and the short rate in JPY over 1949–2003 to justify this model. They obtain some interesting results:

- (1) They show that no-arbitrage prices for zero-coupon bonds are calculated as discounted expected values under the real-world probability measure.
- (2) The model quadratic variation of the discounted GOP fits the empirical values for 1949–2003 rather well.
- (3) They consider two specific models for the shadow short rate Ψ_t for which $r_t = \max(\Psi_t, 0)$, as in the Black model (Black 1995). They provide formulas for zero-coupon bonds and some interest rate derivatives.

An interesting feature: “For the proposed model, an equivalent risk neutral martingale measure is neither possible nor required. Hence we use the benchmark approach where the growth optimal portfolio is chosen as numeraire. Fair derivative prices are then calculated via conditional expectations under the real world probability measure.”

This is a long and complex paper, but it is a promising approach for modeling a low-interest-rate environment.

Akahori, J., and T. Tsuchiya. 2006. What Is the Natural Scale for a Lévy Process in Modelling Term Structure of Interest Rates? *Asia-Pacific Financial Markets* 13(4): 299–313.

In this necessarily mathematical paper the authors show that, under reasonable assumptions, a term structure model based on a Lévy process is arbitrage-free. The authors call these Lévy density term structure models (LDTSM). They show that the LDTSM class includes quadratic term structure models and Gaussian Heath-Jarrow-Morton models. This is a significant theoretical development, but there are no practical examples. There is no discussion of estimating these models.

Diebold, F. X., and C. Li. 2006. Forecasting the Term Structure of Government Bond Yields. *Journal of Econometrics* 130: 337–364.

The authors use the Nelson-Siegel model yield curve as the basis for forecasting interest rates. The yield curve at time t is given as:

$$y_t(\tau) = \beta_{1t} + \beta_{2t} \left(\frac{1 - e^{-\lambda_t \tau}}{\lambda_t \tau} \right) + \beta_{3t} \left(\frac{1 - e^{-\lambda_t \tau}}{\lambda_t \tau} - e^{-\lambda_t \tau} \right)$$

They obtain yields corresponding to end-of-month prices for U.S. Treasury bonds from January 1985 through December 2000. The functions β_{1t} , β_{2t} , and β_{3t} are latent dynamic factors. The authors describe them as follows: The β_{1t} function is a long-term factor that governs the level of the yield curve. The β_{2t} function is a short-term factor that governs the slope, and β_{3t} is a long-term factor that governs the curvature. The function λ_t is set equal to a constant, $\lambda_t = 0.0609$, which maximizes the long-term factor at $\tau = 30$ months; it is constant at this value throughout the paper.

The authors model the factors as a univariate AR(1) processes, thereby creating a dynamic version of the Nelson-Siegel model. They investigate several alternatives and provide statistical test results. The one-month-ahead forecasts are no better than alternatives, but the one-year-ahead are "much superior." They comment on the fact that this model is not arbitrage-free, saying it is not obvious that an arbitrage-free model is necessary or desirable for forecasting.

Hanes, C. 2006. The Liquidity Trap and U.S. Interest Rates in the 1930s. *Journal of Money, Credit, and Banking* 38(1): 163–194.

The author shows that bond yields were negatively related to nonborrowed reserve supply in the U.S. 1930s period of zero overnight interest rates. Bond yields were unrelated to changes in required reserves apart from news about monetary policy. The results are relevant to the debates about monetary policy under low inflation and contradict the view that the interest rate channel of monetary policy is blocked when short-term rates are zero (the so-called liquidity trap).

Realdon, M. 2009. Extended Black Term Structure Models. *International Review of Financial Analysis* 18: 232–239.

The author examines "extended Black" term structure models (EBTSM), which are multifactor extensions of Black's (1995) interest rate model. There are no formulas, but the author shows that EBTSM bond pricing is reasonable as the finite difference numerical solution of a partial differential equation. From the author's conclusion: "EBTSM merit attention, even though they are less tractable than affine term structure models (ATSM) and quadratic term structure models (QTSM)."

Le, A., K. J. Singleton and Q. Dai. 2010. Discrete-Time Affine^Q Term Structure Models with Generalized Market Prices of Risk. *Review of Financial Studies* 23(5): 2184–2227.

The authors study dynamic term structure models (DTCMs) and non-negative DTSMs (NDTSMs) because of the recent experience following the global financial crisis when even US interest rates dropped to zero effectively. The authors develop the formulation and empirical goodness-of-fit of multi-factor DTSMs that enforce a zero lower bound on bond yields.

Bandara, W., and R. Munclinger. 2012. Term Structure Modeling with Structural Breaks: A Simple Arbitrage-Free Approach Available at SSRN: <http://ssrn.com/abstract=1974033>.

The authors start with a general continuous regime-switching model of term structure of interest rates that leads to a general class of affine models. This is an extension of what they call the dynamic Nelson and Siegel model.

Kikuchi, K. 2012. Design and Estimation of a Quadratic Term Structure Model With a Mixture of Normal Distributions. Tokyo: Institute for Monetary and Economic Studies.

The author presents the quadratic mixture of Gaussian term structure model (QMGTM) and derives bond prices using a log-linear approximation. This model does not allow negative yields, considered a drawback of the affine Gaussian term structure model (AGTM). And it captures the fat-tailed feature of changes in yields, which the quadratic Gaussian term structure model (QGTM) fails to do. Estimation requires a nonlinear filtering method. The usual Kalman filter approach does not work. The author estimates and compares the three models—AGTM, QGTM and QMGTM—using monthly data from January 1996 through December 2010 on Japanese government bonds.

Kim, D. H., and K. J. Singleton. 2012. Term Structure Models and the Zero Bound: An Empirical Investigation of Japanese Yields. *Journal of Econometrics* 170: 32–49.

The authors study dynamic term structure models (DTCMs) and nonnegative DTSMs (NDTSMs) because of the recent experience following the global financial crisis when even U.S. interest rates dropped to zero effectively. "This paper

explores in depth the formulation and empirical goodness-of-fit of multi-factor DTSMs that enforce a zero lower bound on bond yields."

They spend some time on the "best known" NDTSM, the Cox-Ingersoll-Ross model. They point out that in the CIR model, if the short rate is zero, then the n factors are also zero, and consequently yields on bonds of all maturities are constant. This is inconsistent with the Japan bond market experience. They show how the CIR model can be generalized by allowing (nonzero) correlation among the factors.

The also consider quadratic Gaussian models, finding that even if the short rate is zero, the long-term yields can fluctuate, thereby allowing the possibility of fitting the Japanese data. They also discuss shadow rate models and their limitations.

The main contribution is estimation of five two-factor NDTSMs, including two affine models, a Gaussian quadratic model and two shadow rate models. Their empirical findings suggest that all of these capture variation in bond yields and risk premiums when the interest rate is near zero. The affine models do not fit the yield distributions as well as the others. Overall the shadow rate models outperform the other models, providing the best fits at both the short and long ends of the yield curve.

Gourieroux, C., and A. Monfort. 2013. Linear-Price Term Structure Models. *Journal of Empirical Finance* 24: 24–41.

The authors present an alternative to the affine term structure models (ATSMs) based on bond prices rather than yields, called the linear term structure model (LTSM). The new approach has interesting technical and practical features. The main feature is that it accounts for "regimes hidden from investors." The authors show that two classes of LTSM provide "quasi explicit" formulas for bond prices and derivatives.

Xiang, J., and X. Zhu. 2013. A Regime-Switching Nelson-Siegel Term Structure Model and Interest Rate Forecasts. *Journal of Financial Econometrics* 11(5): 522–555.

This paper develops and estimates a dynamic Nelson-Siegel model with regime shifts. The authors begin with Diebold and Li's (2006) reinterpretation of the original Nelson-Siegel model (with different notation). In this paper, the yield curve equation is as follows:

$$i_t(\tau) = L_t + S_t \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} \right) + C_t \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right)$$

In the Diebold and Li model, the level factor is β_{1t} , but here it is L_t . In the same way, the curvature factor is $\beta_{3t} = C_t$, and the slope factor is $\beta_{2t} = S_t$.

This paper uses a state-space representation that allows modeling the entire yield curve. This approach does not require ex ante that λ be fixed. A Markov-switching vector autoregressive process governs the dynamic behavior of the three factors. The result is called the regime-switching dynamic Nelson-Siegel model. It allows for one, two or three regimes. The authors estimate the model and do extensive testing. They find that the model fits the yield curve very well. It provides accurate out-of-sample forecasts even at short horizons.

Wu, J. C., and F. D. Xia. 2014. Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound. Cambridge, MA: National Bureau of Economic Research.

The authors propose a new interest rate model using the shadow rate similar to Black's (1995) model. It is an approximation to a nonlinear term structure model designed especially for the situation in which interest rates are near the zero lower bound. They use the model to summarize the macroeconomic effects of unconventional monetary policy at the zero lower bound. It has nice figures (11) and tables (five). An article in the trade press by Zumbrun (2014) is based on this paper.

Zumbrun, J. January 31, 2014. Bernanke Secret Sauce Drops Fed Rate as QE Quantified: Economy. *Bloomberg Business*.

This financial-press article discusses the results of Wu and Xia (2014) on macroeconomic effects of unconventional monetary policy at the zero lower bound on the interest rate. Here is an interesting quote from the article:

The economists, Jing Cynthia Wu and Fan Dora Xia, used a concept known as the “shadow rate” to gauge the impact of quantitative easing and the Fed’s forward guidance on the likely path of interest rates. Their findings: as of December, Fed policy was the equivalent of cutting the benchmark interest rate to minus 1.98 percent, according to Wu at the University of Chicago Booth School of Business and Xia at the University of California at San Diego.

Christensen, J. H. E., and G. D. Rudebusch. 2015. Estimating Shadow-Rate Term Structure Models with Near-Zero Yields. *Journal of Financial Econometrics* 13(2): 226–259.

This is a significant development in interest rate modeling. It is important to recognize the zero lower bound, especially in low-interest-rate environments. The authors extend Black’s approach (1995) based on a shadow interest rate, resulting in a shadow rate version of the Gaussian arbitrage-free Nelson-Siegel model. The authors describe one-, two- and three-factor empirical shadow rate models, based on weekly Japanese government bond yield data over the period from January 6, 1995 to May 3, 2013.

They also compare the shadow rate across the three models. When rates are positive, they find “little disagreement” in the three models. However, when the shadow rate is negative, there can be “pronounced differences.”

CORPORATE BONDS AND CREDIT RATINGS

Jarrow, R. A., et al. 1997. A Markov Model for the Term Structure of Credit Risk Spreads. *Review of Financial Studies* 10(2): 481–523.

This paper presents a model for valuing bonds and derivatives written on the bonds, where the bonds are subject to default. For example, this model applies to corporate bonds, municipal bonds and the associated credit derivatives. The model is presented in both discrete time and continuous time. In either setting, the default time is modeled via a time-homogenous Markov chain with a finite state space representing possible credit classes.

The model presented here is an extension of the model presented by Jarrow and Turnbull (1995). It can be used with any term structure model for default-free bonds, such as the Cox et al. (1985) model or the Heath, et al. (1992) model. The authors show how to estimate the model using observable data, and they provide several illustrative examples.

Bangia, A., et al. 2002. Ratings Migration and the Business Cycle, With Application to Credit Portfolio Stress Testing. *Journal of Banking and Finance* 26(2–3): 445–474.

The main results are estimates of transition matrices for credit ratings. The authors find distinct differences between U.S. expansion and contraction matrices. That is, estimates in an expanding and contracting economy are different. In addition to the transition of ratings, the authors also estimate the quarterly regime-switching matrices, where the two regimes are expansion and contraction. The authors use National Bureau of Economic Research data for the regime-switching estimates and Standard & Poor’s data for the rating transition estimates, both over the period 1981–1998. The authors claim that first-order Markov rating dynamics (as in Jarrow et al. 1997) provides a reasonable practical approximation to actual rating behavior, so long as different transition matrices are used for expansions and contractions.

Parnes, D. 2007. Time Series Patterns in Credit Ratings. *Finance Research Letters* 4(4): 217–226.

This paper provides an alternative to the homogeneous Markov chain dynamics of credit ratings. Here the dynamics is specified as an AR(1) time series for each entry in the transition matrix. The author shows how to estimate the model using S&P credit ratings for 1985–2004.

The paper provides a comparative analysis of the homogeneous Markov chain approach and the alternative internal correlations model. The two techniques were calibrated on data for 1985–1994. Then both models were run for 1995–2004 and compared with actual credit default rates. It turns out that the Markov chain approach overestimates transition probabilities and projects much higher default rates than observed. The proposed alternative provides a better fit (but still overestimates).

Acharya, V. V., et al. 2013. Liquidity Risk of Corporate Bond Returns: Conditional Approach. *Journal of Financial Economics* 110(2): 358–386.

The paper uses two liquidity risk factors. The stock illiquidity index is the market's average price-impact measure. It is calculated as the equally weighted average of the daily ratio of absolute stock return to its daily dollar volume and averaged over the days of the month to provide a monthly illiquidity measure using NYSE and Amex stocks. The bond illiquidity measure is the equally weighted quoted bid-ask spread on on-the-run short-maturity treasuries. The authors show that liquidity risk in the bond market is conditional on the state of the economy. Liquidity risk is more important in times of financial distress. The authors are able to predict out-of-sample bond returns for the stress years 2008–2009.

STOCK PRICE MODELS

Black, F., and M. Scholes. 1973. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy* 81(3): 637–654.

This is the paper in which Black and Scholes published the famous equation, now called the Black-Scholes-Merton equation, for valuing European-style call and put options on a stock. Merton published about the same time as Merton (1973). Merton and Scholes received the 1997 Nobel Memorial Prize in Economic Sciences for their work. Though ineligible for the prize because of his death in 1995, Black was mentioned as a contributor by the Swedish Academy.

Options on stocks or stock indices arise in insurance models as, for example, return guarantees in variable annuities, unit-linked life insurance, and universal life insurance.

Boyle, P. P. 1977. Options: A Monte Carlo Approach. *Journal of Financial Economics* 4(3): 323–338.

This is one of the first papers on option valuation via simulation. Boyle shows how to value European-style options on dividend-paying stocks. This is seminal work on Monte Carlo-based option pricing. This paper facilitated the 1980s explosion in the world of derivatives.

Tilley, J. A. 1993. Valuing American Options in a Path Simulation Model. *Transactions of the Society of Actuaries* 46: 499–549.

This paper demonstrates that American-style options can be valued efficiently in a simulation model. It removes what has been considered a major impediment to the use of simulation models for valuing financial instruments or insurance company asset liability valuation. This is a general algorithm for estimating the value of American options on an underlying security for which the arbitrage-free probability distribution of paths through time can be simulated. The general algorithm is tested on a put option on a non-dividend-paying stock for which the exact option premium can be determined.

Amaral, L. A. N., et al. 2000. The Distribution of Returns of Stock Prices. *International Journal of Theoretical and Applied Finance* 3(3): 365–369.

This is an empirical study of stock returns. The stock price data comes from the trades and quotes database for 1994–1995 and the Center for Research and Security Prices database for 1962–1996. There are three findings:

- (1) The distribution of normalized returns for individual companies is consistent with a power law behavior characterized by an exponent $\alpha \approx 3$.
- (2) The distributions of returns retain the same functional form for a wide range of time scales Δt , varying over three orders of magnitude, $5 \leq \Delta t \leq 6,240$ minutes = 16 days.
- (3) For $\Delta t > 16$ days, the distribution of returns appears to slowly converge to a Gaussian distribution.

Hartmann, P., et al. 2004. Asset Market Linkages in Crisis Periods. *Review of Economics and Statistics* 86(1): 313–326.

The authors apply an extreme dependence measure to weekly returns of stocks and bonds in the G-5 countries during 1987–1999. There are some important results that should be considered when modeling stock and bond returns in more than one country. Extreme losses are much higher for stock indices than government bond indices. Stock index extremes are asymmetric, in that the absolute extreme loss in stock index returns exceeds the maximum positive

return. Stock markets have co-crashes and flight-to-quality effects during periods of market turbulence. Stock market crashes on the order of a 20% weekly loss and a bond market weekly loss of 8% are rare events, occurring once or twice in a lifetime. Stock market co-crashes occur in approximately one out of five to eight crashes. The number is lower for bond markets and even less likely for stock and bond co-crashes. A widespread securities meltdown in the main industrialized countries happens much less often than once in 200 years. It is nevertheless much higher than a normal-distribution-based analysis would lead one to believe. One implication is that a severe market crisis spilling over to another market cannot be a prevalent phenomenon among G-5 countries.

Hilal, S., et al. 2011. Hedging the Black Swan: Conditional Heteroskedasticity and Tail Dependence in S&P500 and VIX. *Journal of Banking and Finance* 35(9): 2374–2387.

Kristensen, D., and A. Mele. 2011. Adding and Subtracting Black-Scholes: A New Approach to Approximating Derivative Prices in Continuous-Time Models. *Journal of Financial Economics* 102(2): 390–415.

This paper provides a method of approximating derivative prices in multifactor continuous-time models. For a model with no closed-form solution, it requires a closely related model that has a closed-form solution. A special case is given in the examples. The difference of the solutions is approximated by a finite sum of differences related to the partial differential equation of the difference in the prices. The authors illustrate this with applications to a generalized Black-Scholes option model with stochastic volatility, term structure of interest rate models including CIR and Vasicek models, and stock price models with jumps. They provide many examples showing how to apply this method.

Grundke, P., and S. Polle. 2012. Crisis and Risk Dependencies. *European Journal of Operational Research* 223(2): 518–528.

The authors set out to answer two questions.

- Is it possible to describe the multivariate stochastic dependence between the returns of different asset classes by some standard parametric families of copulas?
- If so, how did the stochastic dependence between the returns of different asset classes change during the recent financial crisis?

They use data on six assets classes (credit, stocks, bonds, currencies, commodities and real estate) during January 2, 2006, to October 21, 2008. The authors report “rather bad news” because of the observed weakness of statistical goodness-of-fit tests in identifying the correct copula. How large the consequences of this model risk are depends on the application. Whether a specific copula can be rejected or not depends on the asset class and the time period. Different tests for copulas can yield very different results. Finally, they suggest some alternatives to goodness-of-fit tests of copulas, such as a graphical tool or Akaike’s information.

Andersen, L., and A. Lipton. 2013. Asymptotics for Exponential Lévy Processes and Their Volatility Smile: Survey and New Results. *International Journal of Theoretical and Applied Finance* 16(1): 1–91.

This is an unusual paper because it has introductory material on foreign-currency exchange trading and Lévy processes. It is unusually long, almost 100 pages. It has a good introduction to foreign-exchange markets, background for Lévy processes, and discussion of the Lewis-Lipton option price formula. The main contribution is three sections on asymptotics. This is from the paper’s abstract: “Our main conclusion is that studying asymptotic properties of the implied volatility, while theoretically exciting, is not always practically useful because the domain of validity of many asymptotic expressions is small.”

Bollerslev, T., et al. 2013. Jump Tails, Extreme Dependencies, and the Distribution of Stock Returns. *Journal of Econometrics* 172: 307–324.

The authors apply new statistical techniques for analyzing jumps in individual stock returns. For a large panel of 50 stocks and the S&P 500, they find that idiosyncratic and systematic jumps are both heavy tailed. They also find asymptotic tail dependence between individual stocks and the index.

Durham, G. B. 2013. Risk-Neutral Modeling with Affine and Non-affine Models. *Journal of Financial Econometrics* 11(4): 650–681.

This is an important development in asset-pricing models, although necessarily long and complex. The author considers seven log volatility models and seven affine models classified as follows: no jumps, jumps in returns alone, and jumps in both returns and volatility.

A given model is assessed in these steps:

- Given observed option prices, invert the risk-neutral measure to get the implied volatility.
- Use the option data and implied volatility to compute the log-likelihood.
- Examine some diagnostics. For example, the generalized residuals u_t should be independent and identically distributed uniform on $(0, 1)$, and their transforms $\tilde{u}_t = \Phi^{-1}(u_t)$ should be identically distributed standard normal variables. Diagnostics such as normal-quintile plots are applied in this step.

The author finds that log volatility models are “dramatically” better than affine models. Affine models have problems: All affine models exhibit substantial autocorrelation in squared volatility residuals. The square root specification for volatility does not fit the data. Also, affine models with exponential jumps are able to match the right tail but not the left. A final quote from the paper:

All the models are rejected at far beyond conventional significance levels on at least one test. One would have to be almost hopelessly optimistic to believe that any of the models examined in this article was the true data-generating process. Failure to reject a model in exercises such as this should more likely be interpreted as a sign of insufficient sample size or tests with low power rather than an indication that one has found the true data-generating process. Powerful diagnostics are a good thing; failure to find evidence of defective models is a serious liability.

Maheua, J. M., et al. 2013. Do Jumps Contribute to the Dynamics of the Equity Premium? *Journal of Financial Economics* 110(2): 457–477.

This paper shows the importance of jumps in equity models. Jump risk is priced in the market index, and it contributes to the equity premium. The authors show that it is possible to identify jump risk and its contribution to the market risk-return trade-off using equity data only. It is not necessary to use options data.

Wang, Y.-C., et al. 2013. A Revisit to the Dependence Structure Between the Stock and Foreign Exchange Markets: A Dependence-Switching Copula Approach. *Journal of Banking and Finance* 37(5): 1706–1719.

This paper provides a dependence-switching copula model to describe the dependence structure between stock markets and foreign-exchange markets. In this way, it allows for dependence of bull and bear markets associated with currency appreciation and depreciation. This is important in an international economic scenario generator.

The authors estimate the model based on stock indices for Canada, France, Germany, Italy, Japan and the United Kingdom and currency exchange rates expressed in U.S. dollars. In this setting, they find dependence and tail dependence are asymmetric in a negative-correlation regime but symmetric in a positive-correlation regime. Exchange rate exposure effects or portfolio-rebalancing effects dominate for most countries on most occasions.

INVESTMENT MODELS, ECONOMIC SCENARIO GENERATORS

Redington, F. M. 1952. Review of the Principles of Life-Office Valuations. *Journal of the Institute of Actuaries (1886–1994)* 78(3): 286–340.

The main contribution is the rigorous exposition of asset-liability matching (ALM). The author has in mind application to life insurance, but the development is rather general with applications to life office management. Although the author uses the terms "mean" and "spread about the mean," he does not define random variables explicitly. This is fundamental work on ALM.

Wilkie, A. D. 1986. A Stochastic Investment Model for Actuarial Use. *Transactions of the Faculty of Actuaries* 39: 341–403.

This is one of the first papers on stochastic investment models for actuarial applications. The model has four variables: a retail price index $Q(t)$, an index of stock dividends $D(t)$, dividend yield $Y(t)$, and a default-free interest

rate $C(t)$. The variables are modeled in a cascade structure.

The author describes the model estimation, sensitivity analysis and applications. The suggested applications include estimation of reserves for unit-linked life insurance, the estimation of cost of guarantees for indexed annuities, portfolio selection, and estimation of premiums.

Wilkie, A. D. 1992. Stochastic Investment Model for the 21st Century Actuaries. *Transactions of the 24th International Congress of Actuaries* 5: 119–137.

Wilkie reviews the history of computer-based models in actuarial work, particularly in the United Kingdom. In Section 2, he describes the historical development of stochastic investment models. An early example was the Institute of Actuaries Maturity Guarantees Working Party approach, which was developed later by the author. Wilkie mentions models developed by Tilley (1990) and (1992), but those are bond-yield-only models. In Section 3, Wilkie reviews applications of stochastic investment models. The original MGWP application was to investigate contingency reserves for unit-linked life insurance.

Wilkie cites many applications in the 1980s and 1990s to life insurance, general (property-casualty) insurance and pension plans. In Section 5, Wilkie briefly discusses generalizations, and in Section 6, he presents an extension of the model to several countries. This requires the fundamental variables—price index $Q(t)$, dividend yield $Y(t)$, dividend index $D(t)$ and interest rate $C(t)$ —for each country, as well as the exchange rates between countries and the interaction of all the variables. Wilkie describes and estimates the international model for three countries: the United Kingdom in 1923–1990, the United States in 1926–1989 and France in 1951–1989.

Ahlgrim, K. C., et al. 2005. Modeling Financial Scenarios: A Framework for the Actuarial Profession. *Proceedings of the Casualty Actuarial Society* 177(2): 177–238.

In May 2001, the Casualty Actuarial Society (CAS) and the Society of Actuaries (SOA) jointly issued a request for proposals on the research topic “Modeling of Economic Series Coordinated With Interest Rate Scenarios.” This paper summarizes the authors’ research project initiated in response to the joint request for proposals. The authors’ project involves the construction of a financial scenario model that simulates a variety of economic variables over a 50-year period. The variables projected by this model include interest rates, inflation, equity returns, dividend yields, real estate returns and unemployment rates for the United States.

The paper includes a good discussion of modeling issues and a review of the literature. The third section describes the component models. The authors also provide a discussion of the data used to calibrate the model. There is also a discussion of simulated economic scenarios generated with the model, including correlations among the variables.

Peters, M. 2007. Economic Scenarios and Economic Capital. Enterprise Risk Management Symposium.

The author presented Allstate’s ESG at the Enterprise Risk Management Symposium in 2007. Allstate is a large insurance company with 13 lines of business, including life and property-casualty. According to the author, the model covers Allstate’s core operations in life and property-casualty. Allstate uses a cascade model illustrated in the cited presentation.

Ahlgrim, K. C., et al. 2008. A Comparison of Actuarial Financial Scenario Generators. *Variance: Advancing the Science of Risk* 2(1): 111–134.

This paper is another report of the authors’ research project “Modeling of Economic Series Coordinated with Interest Rate Scenarios.” They include a discussion of the American Academy of Actuaries (AAA) models for equity returns and interest rates. One section is devoted to cash flow testing set up to illustrate the differences between the CAS-SOA model and the AAA C-3 model. One test is based on term life insurance with the net single premium invested in T-bills (50%), large-cap stocks (25%) and small-capstocks (25%). The second example is based on a property-liability loss reserve situation. In this case, a loss reserve of \$10 million is established, which is supported by \$10 million in assets, because property-liability insurance accounting does not permit discounting of loss reserves. The CAS-SOA model generally leads to a wider set of distributions than the AAA model.

McNeil, A. J., and A. D. Smith. 2012. Multivariate Stress Scenarios and Solvency. *Insurance: Mathematics and Economics* 50: 299–308.

From the authors' abstract: "We show how the probabilistic concepts of half-space trimming and depth may be used to define convex scenario sets $Q(\alpha)$ for stress testing the risk factors that affect the solvency of an insurance company over a prescribed time period. By choosing the scenario in $Q(\alpha)$ which minimizes net asset value at the end of the time period, we propose the idea of the least solvent likely event (LSLE) as a solution to the forward stress testing problem. By considering the support function of the convex scenario set $Q(\alpha)$, we establish theoretical properties of the LSLE when financial risk factors can be assumed to have a linear effect on the net assets of an insurer. In particular, we show that the LSLE may be interpreted as a scenario causing a loss equivalent to the Value-at-Risk (VaR) at confidence level α , provided the α -quantile is a sub-additive risk measure on linear combinations of the risk factors. In this case, we also show that the LSLE has an interpretation as a per-unit allocation of capital to the underlying risk factors when the overall capital is determined according to the VaR. These insights allow us to define alternative scenario sets that relate in similar ways to coherent measures, such as expected shortfall. We also introduce the most likely ruin event (MLRE) as a solution to the problem of reverse stress testing."

Bibliography

- Black, F. 1995. Interest Rates as Options. *Journal of Finance* 50(5): 1371–1376.
- Black, F., and P. Karasinski. 1991. Bond and Option Pricing When Short Rates Are Lognormal. *Financial Analysts Journal* 47 (July): 52–59.
- Brennan, M. J., and E. S. Schwartz. 1979. A Continuous-Time Approach to the Pricing of Bonds. *Journal of Banking and Finance* 3: 133–155.
- Cox, J. C., J. E. Ingersoll and S. A. Ross. 1985. A Theory of the Term Structure of Interest Rates. *Econometrica* 53: 385–407.
- Diebold, F. X., and C. Li. 2006. Forecasting the Term Structure of Government Bond Yields. *Journal of Econometrics* 130: 337–364.
- Heath, D., R. Jarrow and A. Morton. 1992. Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation. *Econometrica* 60: 77–105.
- Ho, T. S. Y., and S.-B. Lee. 1986. Term Structure Movements and Pricing Interest Rate Contingent Claims. *Journal of Finance* 41(5): 1011–1029.
- Jarrow, R. A., D. Lando and S. M. Turnbull. 1997. A Markov Model for the Term Structure of Credit Risk Spreads. *Review of Financial Studies* 10(2): 481–523.
- Jarrow, R. A., and S. M. Turnbull. 1995. Pricing Derivatives on Financial Securities Subject to Credit Risk. *Journal of Finance* 50(1): 53–85.
- Merton, R. C. 1973. Rational Theory of Option Pricing. *Bell Journal of Economics and Management Science* 4: 141–183.
- Tilley, J. A. 1990. A Stochastic Yield Curve for Asset/Liability Simulation. *Actuarial Approach for Financial Risks: First AFIR International Colloquium*, Paris, April 23–27, International Actuarial Association.
- . 1992. An Actuarial Layman's Guide to Building Stochastic Interest Rate Generators. *Transactions of the Society of Actuaries* 44: 509–538.
- Vasicek, O. 1977. An Equilibrium Characterisation of the Term Structure. *Journal of Financial Economics* 5(2): 177–188.
- Wu, J. C., and F. D. Xia. 2014. Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound. Cambridge, MA: National Bureau of Economic Research.
- Zumbrun, J. January 31, 2014. Bernanke Secret Sauce Drops Fed Rate as QE Quantified: Economy. *Bloomberg Business*.

Glossary

Arbitrage-Free Model — An arbitrage is defined as an explicit trading strategy involving the simultaneous buying and selling of securities, possibly quite complicated, for which either the trader would receive a positive cash flow today with no further payment obligations during the remainder of the trade horizon, or the trader would enter into a trade at zero net cost and receive only non-negative cash flows during the remainder of the trade horizon, with a positive cash flow occurring with probability greater than zero. The model is said to be arbitrage-free if there do not exist any self-financing trading strategies that produce an arbitrage.

Binomial Interest Rate Model — A discrete-time model of interest rates, where from each current state (interest rate level), there are two possible subsequent states. A binomial model recombines, meaning that the structure of possible states forms a lattice, as opposed to an exponentially growing binary tree.

Black-Box Model — A model where the mathematical underpinnings are not disclosed. In general, a user of a black-box model has access to only the inputs and outputs of the model.

Calibration — The process of setting the parameters of the equations within an ESG model to produce the distributions and dynamics (e.g., volatility, correlations, tail characteristics) of economic and financial variables that are required by the application for which they are being used. The term *calibration* is often used interchangeably with the term *parameterization* in general usage. In this report, we use the term *calibration* to describe the process of setting the parameters of the model equations and distributions to reflect market consistent conditions, and reserve the term *parameterization* to describe the process of setting the coefficients and parameters of the model equations and distributions to reflect real-world behavior.

Cascade Model — A model framework whereby each subsequent variable depends only on prior values of the variable and the values of variables that lie above them in the cascade structure.

Closed-Form Expression — A mathematical expression that can be calculated directly, without requiring iterative methods. Closed-form expressions are usually written in terms of familiar functions or operations.

Deterministic — A quality of a system that means no randomness is involved in developing future states of the system.

Drawdown Risk — The maximum potential loss from peak to trough.

Economic Capital — The amount of capital that a firm, usually in financial services, needs to ensure that the company stays solvent, based on an economic view of firm solvency (i.e., able to fulfill financial obligations).

Economic Capital Models — Models used by financial institutions to determine their economic capital.

Economic Scenario Generator (ESG) — A software tool that simulates future paths of economies and financial markets, and illuminates the nature of risk elements within the economy that drive financial variability.

Expert Judgment — The use of subjective and informed judgment by experienced professionals in the construction, calibration and parameterization of an ESG.

Financial Markets — Markets in which people trade financial securities, options, commodities, currencies and similar items with transparent pricing where market forces of supply and demand determine prices of these items.

Finite Difference Method — A type of numerical method used for solving differential equations by discretizing the equations, approximating derivatives in the equations with finite differences.

Fourier Inversion Method — A specific process of interpreting model output to determine the model source specification.

Fundamental Theorems of Asset Pricing — A collection of theorems that provide necessary and sufficient conditions for a financial market to be arbitrage-free and to be complete (i.e., every asset has one price in every possible state of the world).

Independent and Identically Distributed (IID) — Quality of a set of random variables, that means each random variable in the set having the same probability distribution as the others and all being mutually independent.

Jump Process — A type of stochastic process that has discrete discontinuous movements, called jumps, rather than continuous movements.

Kalman Filter — An algorithm involving consecutive cycles of predicting the state of an observed variable based on a model, comparing that prediction with the realized outcome in the historical observed data, and updating model parameters to achieve optimal predictive power.

Least Square Monte Carlo — A two-step procedure used in valuing American options. The first step is a backward induction process, in which values are assigned to states at every time increment by starting at the final states and working backward through time. At each time step, the value is determined via a least squares regression against market price of the option at that state and point in time. Once every possible state at each time is given a value, the second step of the process is to move forward in time, making an optimal decision on option exercise for each state.

Liquidity — When applied to a financial instrument, the degree to which the asset or security can be traded in the market without affecting the instrument's price. Generally, high liquidity is associated with high trading volume.

Market Consistent — Quality of prices of instruments that are generated within a valuation model that means they are consistent with the prices at which the instruments trade in the market.

Markov Chain — A “memoryless” random process, in which the probability of the next state of the process depends only on the current state, with no regard to preceding states.

Martingale Process — A sequence of random variables where the expected value of the next random variable in the sequence is equal to the value of the current random variable and does not depend on prior random variable values.

Maximum-Likelihood Estimation — A method for estimating statistical model parameters via maximizing a likelihood function, given a data set.

Monte Carlo Process — A broad class of computational algorithms that rely on repeated random sampling to obtain numerical results.

Nested Stochastics (or Stochastic on Stochastic) — A simulation approach where variables are projected using a real-world calibration (outer loop) and pricing is performed at each future real-world node, using a second stochastic simulation (inner loop) that is based on a market-consistent calibration. This simulation technique is used when pricing of contingent cash flows is required as part of a risk management application.

Ordinary Differential Equations — Equations that involve the ordinary (as opposed to partial) derivatives of functions.

Parameterization — The process of setting the coefficients and parameters of the model equations and distributions to reflect real-world behavior. The term is often used interchangeably with the term *calibration*. In this report, we use the term *calibration* to describe the process of setting the parameters of the model equations and distributions to reflect market-consistent conditions.

Pathwise Behavior — The behavior of a random variable along a given path over time.

Problem Discovery — A model validation approach where a user first runs the model, creating a large output data set, and then tries to discover problems with the output.

Projection Period — The period over which a given variable within a model is projected into the future.

Ratings Migration — Changing of a fixed-income security's credit ranking or credit grade over time.

Real-World Model — A model that attempts to represent the future risk-reward dynamics that may actually occur in the economy and capital markets.

Reduced-Form Models — In credit modeling, a specific kind of simplified model of bond default. In general, reduced-form models are in contrast to structural models that produce solutions by modeling the cause-and-effect relationships; reduced-form models try to capture dynamics using fewer inputs and/or parameters.

Regime-Switching Model — In asset modeling, a stochastic model that utilizes different volatility regimes, with a stochastic process for switching between regimes.

Replicating Portfolio — A portfolio, typically of assets, that replicates the expected cash flows or other properties of another specific portfolio of assets or liabilities. Generally, the replicating portfolio is constructed for market pricing of portfolios with unknown value or efficiently producing representative simulations of more complex portfolios.

Risk-Neutral Model — A model where all investments have the same expected return, regardless of risk.

Scenario — One specific trial from a scenario set.

Scenario Set — The set of all scenarios that make up a stochastic model run.

Simulation — The representation of the operations or features of one process (or system) through the use of another.

State Price Density — A density function over states, in order to price assets under an arbitrage-free system and using real-world probabilities.

Steady-State Behavior — The quality of results from a random process in the case where the future distribution will be consistent with prior-period results.

Stochastic — A system or process involving the use of random variables.

Stochastic Differential Equations — Differential equations involving stochastic processes, for which the solution is also a stochastic process.

Stochastic Volatility Model — A model in which the volatility of a stochastic process is itself modeled by a random process.

Structured Investment Vehicle (SIV) — A pool of assets and liabilities intended to earn a credit spread between longer-term assets in the pool and shorter-term liabilities. The shorter-term liabilities of the SIV generally are leveraged many times over, which may allow for vulnerability to liquidity meltdowns.

Stylized Fact — Generalized interpretations of empirical findings, both qualitative and quantitative, that provide a basis for consistent understanding of markets or economic drivers across a wide range of instruments, markets, and time periods.

Strategic Asset Allocation — A portfolio strategy that involves setting target allocations based on optimizing specific quantities or maintaining specific targets (such as optimizing expected return for a target volatility in the portfolio).

Structural Models — A model for estimating deep “structural” parameters of theoretical economic models. These models are usually contrasted with reduced-form models.

Swaptions — An option to engage in a swap transaction under certain terms; the owner of a swaption is not obligated to enter the specified swap.

Swaption Volatility Surface — A three-dimensional graphical surface of implied swaption volatilities.

Time Horizon — The length of time over which a process is to be modeled.

Validation — A process to ensure that the estimation of an ESG’s parameters results in simulated behavior that is a good representation of the variable or market under consideration.