

Yield Curve Modeling and Forecasting:

The Dynamic Nelson-Siegel Approach

Francis X. Diebold
University of Pennsylvania

Glenn D. Rudebusch
Federal Reserve Bank of San Francisco

April 29, 2012

Contents

Preface	xi
1 Facts, Factors, and Questions	1
1.1 Three Interest Rate Curves	2
1.2 Zero-Coupon Yields	3
1.3 Yield Curve Facts	4
1.4 Yield Curve Factors	7
1.5 Yield Curve Questions	13
1.5.1 Why use factor models for yields?	13
1.5.2 How should bond yield factors and factor loadings be constructed?	15
1.5.3 Is imposition of “no arbitrage” useful?	15
1.5.4 How should term premiums be specified?	16
1.5.5 How are yield factors and macroeconomic variables related?	17
1.6 Onward	20
2 Dynamic Nelson-Siegel	21
2.1 Curve Fitting	21
2.2 Introducing Dynamics	23
2.2.1 Mechanics	23
2.2.2 Interpretation	24
2.2.3 Alternative factorizations	26
2.3 State-Space Representation	27
2.3.1 Measurement and transition	27

2.3.2	Optimal extractions and predictions	29
2.4	Estimation	30
2.4.1	Static Nelson-Siegel in cross section	30
2.4.2	Two-step DNS	30
2.4.3	One-step DNS	32
2.5	Multi-Country Modeling	37
2.5.1	Global yields	37
2.5.2	Country yields	38
2.5.3	State-space representation	38
2.5.4	Discussion	39
2.6	Risk Management	40
2.7	DNS Fit and Forecasting	42
3	Arbitrage-Free Nelson-Siegel	47
3.1	A Two-Factor Warm-Up	49
3.2	The Duffie-Kan Framework	53
3.3	Making DNS Arbitrage-Free	55
3.3.1	The key result	55
3.3.2	The yield-adjustment term	61
3.3.3	On the Bjork-Christensen-Filipović critique and the yield-adjustment term	65
3.4	Workhorse Models	67
3.5	AFNS Restrictions on $A_0(3)$	69
3.6	Estimation	72
3.7	AFNS Fit and Forecasting	75
4	Extensions	81
4.1	Variations on the Basic Theme	81
4.1.1	Bayesian shrinkage	82
4.1.2	Alternative factor loadings	83
4.1.3	Yield spreads	85
4.1.4	Nonlinearities	86
4.2	Additional Yield Factors	88
4.2.1	Four factors: Dynamic Nelson-Siegel-Svensson (DNSS)	89

4.2.2	Five factors: arbitrage-free generalized Nelson-Siegel (AFGNS)	91
4.2.3	Discussion	95
4.3	Stochastic Volatility	95
4.3.1	DNS formulation	96
4.3.2	AFNS formulation	97
4.3.3	Spanned and unspanned volatility	98
4.4	Macroeconomic Fundamentals	99
4.4.1	Two-step DNS approaches	100
4.4.2	One-step DNS approaches	101
4.4.3	The macroeconomy and yield factors	102
4.4.4	On nominal vs. real yields	104
5	Macro-Finance	107
5.1	Macro-Finance Yield Curve Modeling	107
5.2	Macro-Finance and AFNS	111
5.2.1	Inflation expectations and risk	112
5.2.2	Liquidity and interbank lending rates	118
5.3	Evolving Research Directions	122
5.3.1	Unspanned macroeconomic risks	123
5.3.2	The zero lower bound	123
5.3.3	Bond supply and the risk premium	124
6	Epilogue	127
6.1	Is Imposition of No-Arbitrage Helpful?	129
6.2	Is AFNS the Only Tractable $A_0(3)$ Model?	131
6.3	Is AFNS Special?	132
6.3.1	AFNS' simple structure facilitates special- izations, extensions, and varied uses	132
6.3.2	AFNS has strong approximation-theoretic motivation	133
6.3.3	AFNS restrictions are not rejected	134
	Appendices	137

A Two-Factor AFNS Calculations	137
A.1 Risk-Neutral Probability	137
A.2 Euler Equation	140
B Details of AFNS Restrictions	143
B.1 Independent-Factor AFNS	145
B.2 Correlated-Factor AFNS	147
C The AFGNS Yield Adjustment Term	151
Bibliography	155

List of Figures

1.1	Bond Yields in Three Dimensions	6
1.2	Bond Yields in Two Dimensions	9
1.3	Bond Yield Principal Components	10
1.4	Empirical Level, Slope, and Curvature, and First Three Principal Components, of Bond Yields . . .	12
2.1	DNS Factor Loadings	25
2.2	Out-of-Sample Forecasting Performance: DNS vs. Random Walk	45
4.1	DNSS Factor Loadings	90
4.2	DGNS Factor Loadings	92
5.1	Nominal and Real Yields and BEI Rates	116
5.2	BEI Rates and Expected Inflation	117
5.3	Probabilities of Nonpositive Net Inflation	119
5.4	LIBOR Spreads	122

List of Tables

1.1	Bond Yield Statistics	5
1.2	Yield Spread Statistics	7
1.3	Yield Principal Components Statistics	11
3.1	AFNS Parameter Restrictions on the Canonical $A_0(3)$ Model	72
3.2	Out-of-Sample Forecasting Performance: Four DNS and AFNS Models	76
3.3	Out-of-Sample Forecasting Performance: Random Walk, $A_0(3)$, and $AFNS_{indep}$	78

To our wives

Preface

Understanding the dynamic evolution of the yield curve is important for many tasks, including pricing financial assets and their derivatives, managing financial risk, allocating portfolios, structuring fiscal debt, conducting monetary policy, and valuing capital goods. To investigate yield curve dynamics, researchers have produced a huge literature with a wide variety of models. In our view it would be neither interesting nor desirable to produce an extensive survey. Indeed our desire is precisely the opposite: we have worked hard to preserve the sharp focus of our Econometric Institute and Tinbergen Institute (EITI) Lectures, delivered in Rotterdam in June 2010, on which this book is based.

Our sharp focus is driven by an important observation: most yield curve models tend to be either theoretically rigorous but empirically disappointing, or empirically successful but theoretically lacking. In contrast, we emphasize in this book two intimately-related extensions of the classic yield curve model of Nelson and Siegel (1987). The first is a dynamized version, which we call “dynamic Nelson-Siegel” (DNS). The second takes DNS and makes it arbitrage-free; we call it “arbitrage-free Nelson Siegel” (AFNS). Indeed the two models are just slightly different implementations of a single, unified approach to dynamic yield curve modeling and forecasting. DNS has been highly successful empirically and can easily be made arbitrage-free (i.e.,

converted to AFNS) if and when that is desirable.

Our intended audience is all those concerned with bond markets and their links to the macroeconomy, whether researchers, practitioners or students. It spans academic economics and finance, central banks and NGOs, government, and industry. Our methods are of special relevance for those interested in asset pricing, portfolio allocation, and risk management.

We use this book, just as we used the EITI Lectures, as an opportunity to step back from the signposts of individual journal articles and assess the broader landscape – where we’ve been, where we are, and where we’re going as regards the whats and whys and hows of yield curve modeling, all through a DNS lens. Our methods and framework have strong grounding in the best of the past, yet simultaneously they are very much intertwined with the current research frontier and actively helping to push it outward.

We begin with an overview of yield curve “facts” and quickly move to the key fact: Beneath the high-dimensional set of observed yields, and guiding their evolution, is a much lower-dimensional set of yield factors. We then motivate DNS as a powerful approximation to that dynamic factor structure. We treat DNS yield curve modeling in a variety of contexts, emphasizing both descriptive aspects (in-sample fit, out-of-sample forecasting, etc.) and efficient-markets aspects (imposition of absence of arbitrage, whether and where one would *want* to impose absence of arbitrage, etc.). We devote special attention to the links between the yield curve and macroeconomic fundamentals.

We are pleased to have participated in the DNS research program with talented co-authors who have taught us much *en route*: Boragan Aruoba, Lei Ji, Canlin Li, Jens Christensen, Jose Lopez, Monika Piazzesi, Eric Swanson, Tao Wu, and Vivian Yue. Christensen’s influence, in particular, runs throughout this book.

We are exceptionally indebted to Herman van Dijk and Dick

van Dijk for their intellectual leadership in organizing the EITI Lectures. We are similarly indebted to the team at Princeton University Press, especially Seth Ditchik, for meticulous and efficient administration and production.

We are grateful to many colleagues and institutions for helpful input at various stages. In particular, we thank – without implicating in any way – Caio Almeida, Boragan Aruoba, Jens Christensen, Dick van Dijk, Herman van Dijk, Greg Duffee, Darrell Duffie, Jesús Fernández-Villaverde, Mike Gibbons, Jim Hamilton, Jian Hua, Lawrence Klein, Siem Jan Koopman, Leo Krippner, Jose Lopez, Andre Lucas, Emanuel Mönch, James Morley, Charles Nelson, Ken Singleton, Dongho Song, Jim Steeley, Chuck Whiteman, and Tao Wu. Several anonymous reviewers of the manuscript also provided insightful and valuable comments. For research assistance we thank Fei Chen, Jian Hua, and Eric Johnson. For financial support we are grateful to the National Science Foundation, the Wharton Financial Institutions Center, and the Guggenheim Foundation.

We hope that the book conveys a feeling for the excitement of the rapidly evolving field of yield curve modeling. That rapid evolution is related to, but no excuse for, the many errors of commission and omission that surely remain, for which we apologize in advance.

Francis X. Diebold
Philadelphia 2011
fdiebold@sas.upenn.edu
<http://www.ssc.upenn.edu/~fdiebold>

Glenn D. Rudebusch
San Francisco 2011
glenn.rudebusch@sfrb.org
<http://www.frbsf.org/economics/economists/grudebusch>

Additional Acknowledgment

This book draws on certain of our earlier-published papers, including:

Christensen, J.H.E., Diebold, F.X. and Rudebusch, G.D. (2009), “An Arbitrage-Free Generalized Nelson-Siegel Term Structure Model,” *The Econometrics Journal*, 12, 33-64.

Christensen, J.H.E., Diebold, F.X. and Rudebusch, G.D. (2011), “The Affine Arbitrage-Free Class of Nelson-Siegel Term Structure Models,” *Journal of Econometrics*, 164, 4-20.

Christensen, J.H.E., Lopez, J.A. and Rudebusch, G.D. (2009), “Do Central Bank Liquidity Facilities Affect Interbank Lending Rates?” Unpublished manuscript, Federal Reserve Bank of San Francisco.

Christensen, J.H.E., Lopez, J.A. and Rudebusch, G.D. (2010), “Inflation Expectations and Risk Premiums in an Arbitrage-Free Model of Nominal and Real Bond Yields,” *Journal of Money, Credit, and Banking*, 42, 143-178.

Diebold, F.X., Ji, L. and Li, C. (2006), “A Three-Factor Yield Curve Model: Non-Affine Structure, Systematic Risk Sources, and Generalized Duration,” in L.R. Klein (ed.), *Macroeconomics, Finance and Econometrics: Essays in Memory of Albert Ando*, 240-274. Cheltenham, U.K.: Edward Elgar.

Diebold, F.X. and Li, C. (2006), “Forecasting the Term Structure of Government Bond Yields,” *Journal of Econometrics*, 130, 337-364.

Diebold, F.X., Li, C. and Yue, V. (2008), “Global Yield Curve Dynamics and Interactions: A Generalized Nelson-Siegel Approach,”

Journal of Econometrics, 146, 351-363.

Diebold, F.X., Piazzesi, M. and Rudebusch, G.D. (2005), "Modeling Bond Yields in Finance and Macroeconomics," *American Economic Review*, 95, 415-420.

Diebold, F.X., Rudebusch, G.D. and Aruoba, B. (2006), "The Macroeconomy and the Yield Curve: A Dynamic Latent Factor Approach," *Journal of Econometrics*, 131, 309-338.

Rudebusch, G.D. and Swanson, E. (2008), "Examining the Bond Premium Puzzle with a DSGE Model," *Journal of Monetary Economics*, 55, S111-S126.

Rudebusch, G.D. and Swanson, E. (2011), "The Bond Premium in a DSGE Model with Long-Run Real and Nominal Risks," *American Economic Journal: Macroeconomics*, forthcoming.

Rudebusch, G.D., Swanson, E. and Wu, T. (2006), "The Bond Yield 'Conundrum' from a Macro-Finance Perspective," *Monetary and Economic Studies*, 24, 83-128.

Rudebusch, G.D. and Wu, T. (2007), "Accounting for a Shift in Term Structure Behavior with No-Arbitrage and Macro-Finance Models," *Journal of Money, Credit, and Banking*, 39, 395-422.

Rudebusch, G.D. and Wu, T. (2008), "Macro-Finance Model of the Term Structure, Monetary Policy, and the Economy," *Economic Journal*, 118, 906-926.

Chapter 1

Facts, Factors, and Questions

In this chapter we introduce some important conceptual, descriptive, and theoretical considerations regarding nominal government bond yield curves. Conceptually, just what is it that are we trying to measure? How can we best understand many bond yields at many maturities over many years? Descriptively, how do yield curves tend to behave? Can we obtain simple yet accurate dynamic characterizations and forecasts? Theoretically, what governs and restricts yield curve shape and evolution? Can we relate yield curves to macroeconomic fundamentals and central bank behavior?

These multifaceted questions are difficult yet very important. Accordingly, a huge and similarly multifaceted literature attempts to address them. Numerous currents and cross-currents, statistical and economic, flow through the literature. There is no simple linear thought progression, self-contained with each step following logically from that before. Instead the literature is more of a tangled web; hence it is not our intention to produce a “balanced” survey of yield curve modeling, as it is not clear whether that would be helpful or even what it would mean. On the contrary, in this book we slice through the literature in a calculated way, assembling and elaborating on a very particular approach to yield curve modeling. Our approach is simple yet rigorous, simultaneously in close touch with

modern statistical and financial economic thinking, and effective in a variety of situations. But we are getting ahead of ourselves. First we must lay the groundwork.

1.1 Three Interest Rate Curves

Here we fix ideas, establish notation, and elaborate on key concepts by recalling three key theoretical bond market constructs and the relationships among them: the discount curve, the forward rate curve, and the yield curve. Let $P(\tau)$ denote the price of a τ -period discount bond, i.e., the present value of \$1 receivable τ periods ahead. If $y(\tau)$ is its continuously compounded yield to maturity, then by definition

$$P(\tau) = e^{-\tau y(\tau)}. \quad (1.1)$$

Hence the discount curve and yield curve are immediately and fundamentally related. Knowledge of the discount function lets one calculate the yield curve.

The discount curve and the forward rate curve are similarly fundamentally related. In particular, the forward rate curve is defined as

$$f(\tau) = \frac{-P'(\tau)}{P(\tau)}. \quad (1.2)$$

Thus, just as knowledge of the discount function lets one calculate the yield curve, so too does knowledge of the discount function let one calculate the forward rate curve.

Equations (1.1) and (1.2) then imply a relationship between the yield curve and forward rate curve,

$$y(\tau) = \frac{1}{\tau} \int_0^{\tau} f(u) du. \quad (1.3)$$

In particular, the zero-coupon yield is an equally weighted average of forward rates.

The upshot for our purposes is that, because knowledge of any one of $P(\tau)$, $y(\tau)$, and $f(\tau)$ implies knowledge of the other two, the three are effectively interchangeable. Hence with no loss of generality one can choose to work with $P(\tau)$, $y(\tau)$, or $f(\tau)$. In this book,

following much of both academic and industry practice, we will work with the yield curve, $y(\tau)$. But again, the choice is inconsequential in theory.

Complications arise in practice, however, because although we observe prices of traded bonds with various amounts of time to maturity, we do not directly observe yields, let alone the zero-coupon yields at fixed standardized maturities (e.g., six-month, ten-year, ...), with which we will work throughout. Hence we now provide some background on yield construction.

1.2 Zero-Coupon Yields

In practice, yield curves are not observed. Instead, they must be estimated from observed bond prices. Two historically popular approaches to constructing yields proceed by fitting a smooth discount curve and then converting to yields at the relevant maturities using formulas (1.2) and (1.3) above.

The first discount curve approach to yield curve construction is due to McCulloch (1971) and McCulloch (1975), who model the discount curve using polynomial splines.¹ The fitted discount curve, however, diverges at long maturities due to the polynomial structure, and the corresponding yield curve inherits that unfortunate property. Hence such curves can provide poor fits to yields that flatten out with maturity, as emphasized by Shea (1984).

An improved discount curve approach to yield curve construction is due to Vasicek and Fong (1982), who model the discount curve using exponential splines. Their clever use of a negative transformation of maturity, rather than maturity itself, ensures that forward rates and zero-coupon yields converge to a fixed limit as maturity increases. Hence the Vasicek-Fong approach may be more successful at fitting yield curves with flat long ends.

Notwithstanding the progress of Vasicek and Fong (1982), discount curve approaches remain potentially problematic, as the implied forward rates are not necessarily positive. An alternative and popular approach to yield construction is due to Fama and Bliss

¹See also McCulloch and Kwon (1993).

(1987), who construct yields not from an estimated discount curve, but rather from estimated forward rates at the observed maturities. Their method sequentially constructs the forward rates necessary to price successively longer-maturity bonds. Those forward rates are often called “unsmoothed Fama-Bliss” forward rates, and they are transformed to unsmoothed Fama-Bliss yields by appropriate averaging, using formula (1.3) above. The unsmoothed Fama-Bliss yields exactly price the included bonds. Unsmoothed Fama-Bliss yields are often the “raw” yields to which researchers fit empirical yield curves, such as members of the Nelson-Siegel family, about which we will have much to say throughout this book. Such fitting effectively smooths the unsmoothed Fama-Bliss yields.

1.3 Yield Curve Facts

At any time, dozens of different yields may be observed, corresponding to different bond maturities. But yield curves evolve dynamically; hence they have not only a cross-sectional, but also a temporal, dimension.² In this section we address the obvious descriptive question: How do yields tend to behave across different maturities and over time?

The situation at hand is in a sense very simple – modeling and forecasting a time series – but in another sense rather more complex and interesting, as the series to be modeled is in fact a series of *curves*.³ In Figure 1.1 we show the resulting three-dimensional surface for the U.S., with yields shown as a function of maturity, over time. The figure reveals a key yield curve fact: yield curves move a lot, shifting among different shapes: increasing at increasing

²We will be interested in dynamic modeling and forecasting of yield curves, so the temporal dimension is as important as the variation across bond maturity.

³The statistical literature on functional regression deals with sets of curves and is therefore somewhat related to our concerns. See, for example, Ramsay and Silverman (2005) and Ramsay et al. (2009). But the functional regression literature typically does not address dynamics, let alone the many special nuances of yield curve modeling. Hence we are led to rather different approaches.

Table 1.1: Bond Yield Statistics

Maturity (Months)	\bar{y}	$\hat{\sigma}_y$	$\hat{\rho}_y(1)$	$\hat{\rho}_y(12)$
6	4.9	2.1	0.98	0.64
12	5.1	2.1	0.98	0.65
24	5.3	2.1	0.97	0.65
36	5.6	2.0	0.97	0.65
60	5.9	1.9	0.97	0.66
120	6.5	1.8	0.97	0.68

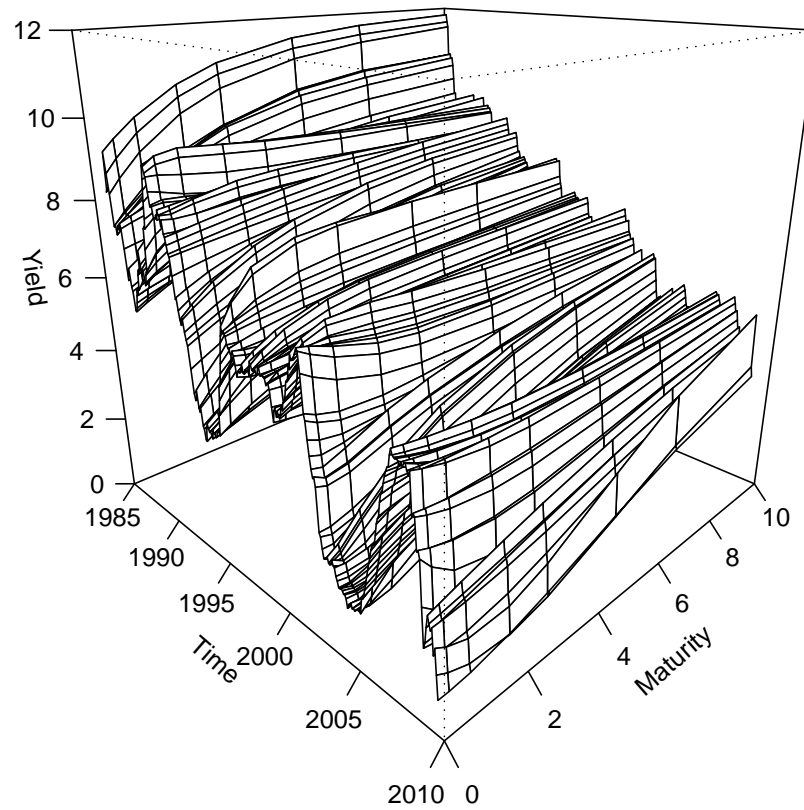
Notes: We present descriptive statistics for end-of-month yields at various maturities. We show sample mean, sample standard deviation, and first- and twelfth-order sample autocorrelations. Data are from the Board of Governors of the Federal Reserve System. The sample period is January 1985 through December 2008.

or decreasing rates, decreasing at increasing or decreasing rates, flat, U-shaped, etc.

In Table 1.1 we show descriptive statistics for yields at various maturities. Several well-known and important yield curve facts emerge. First, time-averaged yields (the "average yield curve") increase with maturity; that is, term premia appear to exist, perhaps due to risk aversion, liquidity preferences, or preferred habitats. Second, yield volatilities decrease with maturity, presumably because long rates involve averages of expected future short rates. Third, yields are highly persistent, as evidenced not only by the very large one-month autocorrelations but also by the sizable twelve-month autocorrelations.

In Table 1.2 we show the same descriptive statistics for yield *spreads* relative to the ten-year bond. Yield spread dynamics contrast rather sharply with those of yield levels; in particular, spreads are noticeably less volatile and less persistent. As with yields, the one-month spread autocorrelations are very large, but they decay

Figure 1.1: Bond Yields in Three Dimensions



Notes: We plot end-of-month U.S. Treasury bill and bond yields at maturities ranging from six months to ten years. Data are from the Board of Governors of the Federal Reserve System, based on Gürkaynak et al. (2007). The sample period is January 1985 through December 2008.

Table 1.2: Yield Spread Statistics

Maturity (Months)	\bar{s}	$\hat{\sigma}_s$	$\hat{\rho}_s(1)$	$\hat{\rho}_s(12)$
6	-1.6	1.3	0.98	0.44
12	-1.4	1.1	0.98	0.46
24	-1.1	0.9	0.97	0.48
36	-0.9	0.7	0.97	0.47
60	-0.6	0.4	0.96	0.44
120	NA	NA	NA	NA

Notes: We present descriptive statistics for end-of-month yield spreads (relative to the ten-year bond) at various maturities. We show sample mean, sample standard deviation, and first- and twelfth-order sample autocorrelations. Data are from the Board of Governors of the Federal Reserve System, based on Gürkaynak et al. (2007). The sample period is January 1985 through December 2008.

more quickly, so that the twelve-month spread autocorrelations are noticeably smaller than those for yields. Indeed many strategies for active bond trading (sometimes successful and sometimes not!) are based on spread reversion.

1.4 Yield Curve Factors

Multivariate models are required for sets of bond yields. An obvious model is a vector autoregression or some close relative. But unrestricted vector autoregressions are profligate parameterizations, wasteful of degrees of freedom. Fortunately, it turns out that financial asset returns typically conform to a certain type of *restricted* vector autoregression, displaying *factor structure*. Factor structure is said to be operative in situations where one sees a high-dimensional object (e.g., a large set of bond yields), but where that high-dimensional object is driven by an underlying lower-dimensional set of objects, or

“factors.” Thus beneath a high-dimensional seemingly complicated set of observations lies a much simpler reality.

Indeed factor structure is ubiquitous in financial markets, financial economic theory, macroeconomic fundamentals, and macroeconomic theory. Campbell et al. (1997), for example, discuss aspects of empirical factor structure in financial markets and theoretical factor structure in financial economic models.⁴ Similarly, Aruoba and Diebold (2010) discuss empirical factor structure in macroeconomic fundamentals, and Diebold and Rudebusch (1996) discuss theoretical factor structure in macroeconomic models.

In particular, factor structure provides a fine description of the term structure of bond yields.⁵ Most early studies involving mostly long rates (e.g., Macaulay (1938)) implicitly adopt a single-factor world view, where the factor is the level (e.g., a long rate). Similarly, early arbitrage-free models like Vasicek (1977) involve only a single factor. But single-factor structure severely limits the scope for interesting term structure dynamics, which rings hollow both in terms of introspection and observation.

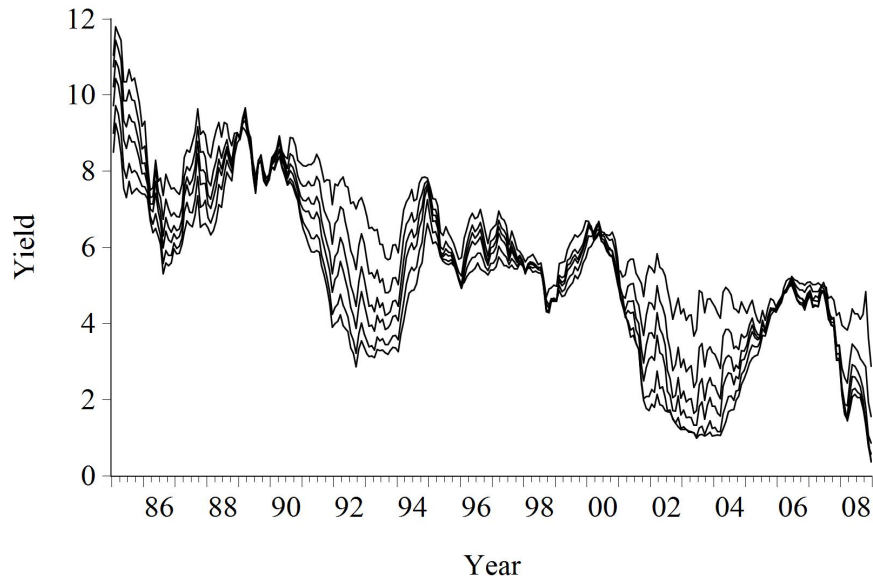
In Figure 1.2 we show a time-series plot of a standard set of bond yields. Clearly they *do* tend to move noticeably together, but at the same time, it’s clear that more than just a common level factor is operative. In the real world, term structure data – and correspondingly, modern empirical term structure models – involve *multiple* factors. This classic recognition traces to Litterman and Scheinkman (1991), Willner (1996) and Bliss (1997), and it is echoed repeatedly in the literature. Joslin et al. (2010), for example, note that:

“The cross-correlations of bond yields are well described by a low-dimensional factor model in the sense that the first three principal components of bond yields...explain well over 95 percent of their variation....Very similar

⁴Interestingly, asset pricing in the factor framework is closely related to asset pricing in the pricing kernel framework, as discussed in Chapter 11 of Singleton (2006).

⁵For now we do not distinguish between government and corporate bond yields. We will consider credit risk spreads later.

Figure 1.2: Bond Yields in Two Dimensions



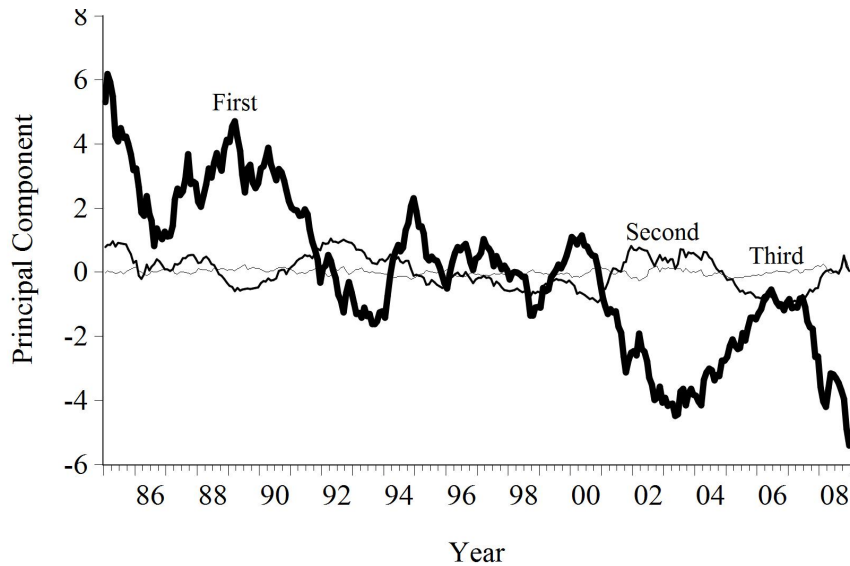
Notes: We plot end-of-month U.S. Treasury bill and bond yields at maturities of 6, 12, 24, 36, 60, and 120 months. Data are from the Board of Governors of the Federal Reserve System, based on Gürkaynak et al. (2007). The sample period is January 1985 through December 2008.

three-factor representations emerge from arbitrage-free, dynamic term structure models...for a wide range of maturities.”

Typically three factors, or principal components, are all that one needs to explain most yield variation. In our data set the first three principal components explain almost one hundred percent of the variation in bond yields; we show them in Figure 1.3 and provide descriptive statistics in Table 1.3.

The first factor is borderline nonstationary. It drifts downward over much of the sample period, as inflation was reduced relative to its high level in the early 1980s. The first factor is the most variable

Figure 1.3: Bond Yield Principal Components



Notes: We show the first, second, and third principal components of bond yields in dark, medium and light shading, respectively.

but also the most predictable, due to its very high persistence. The second factor is also highly persistent and displays a clear business cycle rhythm. The second factor is less variable, less persistent, and less predictable than the level factor. The third factor is the least variable, least persistent, and least predictable.

In Figure 1.4 we plot the three principal components (factors) against standard empirical yield curve level, slope and curvature measures (the 10-year yield, the 10Y-6M spread, and a 6M+10Y-2*5Y butterfly spread, respectively). The figure reveals that the three bond yield factors effectively *are* level, slope, and curvature. This is important, because it implies that the different factors likely have different and specific macroeconomic determinants. Inflation, for example, is clearly related to the yield curve level, and the stage

Table 1.3: Yield Principal Components Statistics

PC	$\hat{\sigma}$	$\hat{\rho}(1)$	$\hat{\rho}(12)$	R^2
First	2.35	0.97	0.67	0.98
Second	0.52	0.97	0.45	0.95
Third	0.10	0.83	0.15	0.70

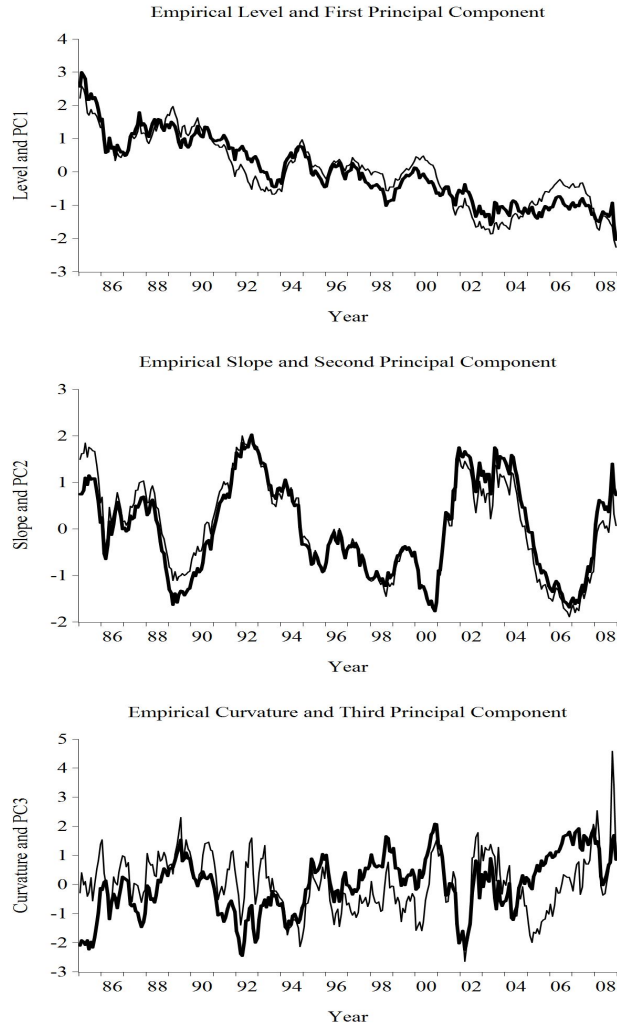
Notes: We present descriptive statistics for the first three principal components of end-of-month U.S. government bill and bond yields at maturities of 6, 12, 24, 36, 60, and 120 months. We show principal component sample standard deviation, first- and twelfth-order principal component sample autocorrelations, and the predictive R^2 (see Diebold and Kilian (2001)) from an $AR(p)$ approximating model with p selected using the Schwartz criterion. Data are from the Board of Governors of the Federal Reserve System, based on Gürkaynak et al. (2007). The sample period is January 1985 through December 2008.

of the business cycle is relevant for the slope. It is also noteworthy that the yield factors are effectively orthogonal due to their exceptionally close links to the principal components, which are orthogonal by construction.

The disproportionate amount of yield variation associated with the common *level* factor, together with its high persistence, explains the broad sweep of earlier-discussed facts, in particular the high persistence of yields and the greatly reduced persistence of yield spreads (because the common level factor vanishes from the spreads). Reality is of course a bit more complicated, as slope and curvature factors are also operative, but the effects of the level factor dominate.

A factor structure for yields with a highly persistent level factor is constrained by economic theory. Economic theory strongly suggests that nominal bond yields should not have unit roots, because the yields are bounded below by zero, whereas unit-root processes have random-walk components and therefore will eventually cross zero almost surely. Nevertheless, the unit root may be a good approximation so long as yields are not too close to zero, as noted by

Figure 1.4: Empirical Level, Slope, and Curvature, and First Three Principal Components, of Bond Yields



Notes: We show the standardized empirical level, slope, and curvature with dark lines, and the first three standardized principal components with lighter lines.

Dungey et al. (2000), Giese (2008) and Jardet et al. (2010), among others.⁶ Work in that tradition, most notably Dungey et al. (2000), finds not only integration but also clear *cointegration*, and the common unit roots associated with cointegration imply factor structure.

1.5 Yield Curve Questions

Thus far we have laid the groundwork for subsequent chapters, touching on aspects of yield definition, data construction, and descriptive statistical properties of yields and yield factors. We have emphasized the high persistence of yields, the lesser persistence of yield spreads, and related, the good empirical approximation afforded by a low-dimensional three-factor structure with highly persistent level and slope factors. Here we roam more widely, in part looking backward, expanding on themes already introduced, and in part looking forward, foreshadowing additional themes that feature prominently in what follows.

1.5.1 Why use factor models for yields?

The first problem faced in term structure modeling is how to summarize the price information at any point in time for the large number of nominal bonds that are traded. Dynamic factor models prove appealing for three key reasons.

First, as emphasized already, factor structure generally provides a highly-accurate empirical description of yield curve data. Because only a small number of systematic risks appear to underlie the pricing of the myriad of tradable financial assets, nearly all bond price information can be summarized with just a few constructed variables or factors. Therefore, yield curve models almost invariably employ a structure that consists of a small set of factors and the associated factor loadings that relate yields of different maturities to those

⁶Alternatively, more sophisticated models, such as the “square-root process” of Cox et al. (1985), can allow for unit-root dynamics while still enforcing yield non-negativity by requiring that the conditional variance of yields approach zero as yields approach zero.

factors.

Second, factor models prove tremendously appealing for statistical reasons. They provide a valuable compression of information, effectively collapsing an intractable high-dimensional modeling situation into a tractable low-dimensional situation. This would be small consolation if the yield data were not well-approximated with factor structure, but again, they are. Hence we're in a most fortunate situation. We *need* low-dimensional factor structure for statistical tractability, and mercifully, the data actually *have* factor structure.

Related, factor structure is consistent with the “parsimony principle,” which we interpret here as the broad insight that imposing restrictions implicitly associated with simple models – even false restrictions that may degrade in-sample fit – often helps to avoid data mining and, related, to produce good out-of-sample forecasts.⁷ For example, an unrestricted vector autoregression provides a very general linear model of yields typically with good in-sample fit, but the large number of estimated coefficients may reduce its value for out-of-sample forecasting.⁸

Last, and not at all least, financial economic *theory* suggests, and routinely invokes, factor structure. We see thousands of financial assets in the markets, but for a variety of reasons we view the risk premiums that separate their expected returns as driven by a much smaller number of components, or *risk factors*. In the equity sphere, for example, the celebrated capital asset pricing model (CAPM) is a single-factor model. Various extensions (e.g., Fama and French (1992)) invoke a few additional factors but remain intentionally very low-dimensional, almost always with less than five factors. Yield curve factor models are a natural bond market parallel.

⁷See Diebold (2007) for additional discussion.

⁸Parsimony, however, is not the only consideration for determining the number of factors needed; the demands of the precise application are of course also relevant. For example, although just a few factors may account for almost all dynamic yield variation and optimize forecast accuracy, more factors may be needed to fit with great accuracy the cross section of yields at a point in time, say, for pricing derivatives.

1.5.2 How should bond yield factors and factor loadings be constructed?

The literature contains a variety of methods for constructing bond yield factors and factor loadings. One approach places structure only on the estimated factors, leaving loadings free. For example, the factors could be the first few principal components, which are restricted to be mutually orthogonal, while the loadings are left unrestricted. Alternatively, the factors could be observed bond portfolios, such as a long-short for slope, a butterfly for curvature, etc.

A second approach, conversely, places structure only on the loadings, leaving factors free. The classic example, which has long been popular among market and central bank practitioners, is the so-called Nelson-Siegel curve, introduced in Nelson and Siegel (1987). As shown by Diebold and Li (2006), a suitably dynamized version of Nelson-Siegel is effectively a dynamic three-factor model of level, slope, and curvature. However, the Nelson-Siegel factors are unobserved, or latent, whereas the associated loadings are restricted by a functional form that imposes smoothness of loadings across maturities, positivity of implied forward rates, and a discount curve that approaches zero with maturity.

A third approach, the no-arbitrage dynamic latent factor model, which is the model of choice in finance, restricts both factors and factor loadings. The most common subclass of such models, affine models in the tradition of Duffie and Kan (1996), postulates linear or affine dynamics for the latent factors and derives the associated restrictions on factor loadings that ensure absence of arbitrage.

1.5.3 Is imposition of “no arbitrage” useful?

The assumption of no arbitrage ensures that, after accounting for risk, the dynamic evolution of yields over time is consistent with the cross-sectional shape of the yield curve at any point in time. This consistency condition is likely to hold, given the existence of deep and well-organized bond markets. Hence one might argue that the real markets are at least approximately arbitrage-free, so that a good yield curve model must display freedom from arbitrage.

But all models are false, and subtleties arise once the inevitability of model misspecification is acknowledged. Freedom from arbitrage is essentially an internal consistency condition. But a misspecified model may be internally consistent (free from arbitrage) yet have little relationship to the real world, and hence forecast poorly, for example. Moreover, imposition of no-arbitrage on a misspecified model may actually *degrade* empirical performance.

Conversely, a model may admit arbitrage yet provide a good approximation to a much more complicated reality, and hence forecast well. Moreover, if reality is arbitrage-free, and if a model provides a very good description of reality, then imposition of no-arbitrage would presumably have little effect. That is, an accurate model would be at least approximately arbitrage-free, even if freedom from arbitrage were not explicitly imposed.

Simultaneously a large literature suggests that coaxing or “shrinking” forecasts in various directions (e.g., reflecting prior views) may improve performance, effectively by producing large reductions in error variance at the cost of only small increases in bias. An obvious benchmark shrinkage direction is toward absence of arbitrage. The key point, however, is that shrinkage methods don’t force absence of arbitrage; rather, they coax things *toward* absence of arbitrage.

If we are generally interested in the questions posed in this subsection’s title, we are also specifically interested in answering them in the dynamic Nelson-Siegel context. A first question is whether our dynamic Nelson-Siegel (DNS) model can be made free from arbitrage. A second question, assuming that DNS can be made arbitrage-free, is whether the associated restrictions on the physical yield dynamics improve forecasting performance.

1.5.4 How should term premiums be specified?

With risk-neutral investors, yields are equal to the average value of expected future short rates (up to Jensen’s inequality terms), and there are no expected excess returns on bonds. In this setting, the expectations hypothesis, which is still a mainstay of much casual and formal macroeconomic analysis, is valid. However, it seems likely

that bonds, which provide an uncertain return, are owned by the same risk-averse investors who also demand a large equity premium as compensation for holding risky stocks. Furthermore, as suggested by many statistical tests in the literature, the risk premiums on nominal bonds appear to vary over time, which suggests time-varying risk, time-varying risk aversion, or both (e.g., Campbell and Shiller (1991), Cochrane and Piazzesi (2005)).⁹

In the finance literature, there are two basic approaches to modeling time-varying term premiums: time-varying quantities of risk or time-varying “prices of risk” (which translate a unit of factor volatility into a term premium). The large literature on stochastic volatility takes the former approach, allowing the variability of yield factors to change over time. In contrast, the canonical Gaussian affine no-arbitrage finance representation (e.g., Ang and Piazzesi (2003)) takes the latter approach, specifying time-varying prices of risk.¹⁰

1.5.5 How are yield factors and macroeconomic variables related?

The modeling of interest rates has long been a prime example of the disconnect between the macro and finance literatures. In the canonical finance model, the short-term interest rate is a linear function of a few unobserved factors. Movements in long-term yields are importantly determined by changes in risk premiums, which also depend on those latent factors. In contrast, in the macro literature, the short-term interest rate is set by the central bank according to its macroeconomic stabilization goals – such as reducing deviations of inflation and output from the central bank’s targets. Furthermore,

⁹However, Diebold et al. (2006b) suggest that the importance of the statistical deviations from the expectations hypothesis may depend on the application.

¹⁰Some recent literature takes an intermediate approach. In a structural dynamic stochastic general equilibrium (DSGE) model, Rudebusch and Swanson (2011) show that technology-type shocks can endogenously generate time-varying prices of risk – namely, conditional heteroskedasticity in the stochastic discount factor – without relying on conditional heteroskedasticity in the driving shocks.

the macro literature commonly views long-term yields as largely determined by expectations of future short-term interest rates, which in turn depend on expectations of the macro variables; that is, possible changes in risk premiums are often ignored, and the expectations hypothesis of the term structure is employed.

Surprisingly, the disparate finance and macro modeling strategies have long been maintained, largely in isolation of each other. Of course, differences between the finance and macro perspectives reflect, in part, different questions, methods, and avenues of exploration. However, it is striking how little interchange or overlap between the two research literatures has occurred in the past. Notably, both the DNS and affine no-arbitrage dynamic latent factor models provide useful statistical descriptions of the yield curve, but in their original, most basic, forms they offer little insight into the nature of the underlying economic forces that drive its movements.

Hence, to illuminate the fundamental determinants of interest rates, researchers have begun to incorporate macroeconomic variables into the DNS and affine no-arbitrage dynamic latent factor yield curve models. For example, Diebold et al. (2006b) provide a macroeconomic interpretation of the DNS representation by combining it with vector-autoregressive dynamics for the macroeconomy. Their maximum likelihood estimation approach extracts three latent factors (essentially level, slope, and curvature) from a set of seventeen yields on U.S. Treasury securities and simultaneously relates these factors to three observable macroeconomic variables (specifically, real activity, inflation, and a monetary policy instrument). By examining the correlations between the DNS yield factors and macroeconomic variables, they find that the level factor is highly correlated with inflation and the slope factor is highly correlated with real activity. The curvature factor appears unrelated to any of the main macroeconomic variables.

The role of macroeconomic variables in a no-arbitrage affine model is explored in several papers. In Ang and Piazzesi (2003), the macroeconomic factors are measures of inflation and real activity, and the joint dynamics of macro factors and additional latent factors are captured by vector autoregressions.¹¹ They find that output shocks

¹¹To avoid relying on specific macro series, Ang and Piazzesi construct

have a significant impact on intermediate yields and curvature, while inflation surprises have large effects on the level of the entire yield curve.

For estimation tractability, Ang and Piazzesi allow only for unidirectional dynamics in their arbitrage-free model; specifically, macro variables help determine yields but not the reverse. In contrast, Diebold et al. (2006b) consider a bidirectional characterization of dynamic macro/yield interactions. They find that the causality from the macroeconomy to yields is indeed significantly stronger than in the reverse direction, but that interactions in both directions can be important. Ang et al. (2007) also allow for bidirectional macro-finance links but impose the no-arbitrage restriction as well, which poses a severe estimation challenge. They find that the amount of yield variation that can be attributed to macro factors depends on whether or not the system allows for bidirectional linkages. When the interactions are constrained to be unidirectional (from macro to yield factors), macro factors can only explain a small portion of the variance of long yields. In contrast, when interactions are allowed to be bidirectional, the system attributes over half of the variance of long yields to macro factors. Similar results in a more robust setting are reported in Bibkov and Chernov (2010).

Finally, Rudebusch and Wu (2008) provide an example of a macro-finance specification that employs more macroeconomic structure and includes both rational expectations and inertial elements. They obtain a good fit to the data with a model that combines an affine no-arbitrage dynamic specification for yields and a small fairly standard macro model, which consists of a monetary policy reaction function, an output Euler equation, and an inflation equation. In their model, the level factor reflects market participants' views about the underlying or medium-term inflation target of the central bank, and the slope factor captures the cyclical response of the central bank, which manipulates the short rate to fulfill its dual mandate to stabilize the real economy and keep inflation close to target. In addition, shocks to the level factor feed back to the real economy through an *ex ante* real interest rate.

their measures of real activity and inflation as the first principal component of a large set of candidate macroeconomic series,

1.6 Onward

In the chapters that follow, we address the issues and questions raised here, and many others. We introduce DNS in chapter 2, we make it arbitrage-free in chapter 3, and we explore a variety of variations and extensions in chapter 4. In chapter 5 we provide in-depth treatment of aspects of the interplay between the yield curve and the macroeconomy. In chapter 6 we highlight aspects of the current frontier, attempting to separate wheat from chaff, pointing the way toward additional progress.