The Dynamic Nelson-Siegel Approach

to

Yield Curve Modeling and Forecasting

Francis X. Diebold
University of Pennsylvania

Glenn D. Rudebusch
Federal Reserve Bank of San Francisco

October 27, 2011
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? 14
    1.5.3 Is imposition of absence of 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

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
CONTENTS

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
CONTENTS

A Two-Factor AFNS Calculations 137
  A.1 Risk-Neutral Probability ....................... 137
  A.2 Euler Equation ................................. 140

B Detail on 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 Yields in Three Dimensions . . . . . . . . . . . . 5
1.2 Yields in Two Dimensions . . . . . . . . . . . . . 9
1.3 Yield Principal Components . . . . . . . . . . . . 10
1.4 Empirical Level, Slope, and Curvature, and First
Three Principal Components . . . . . . . . . . . . 12

2.1 DNS Factor Loadings . . . . . . . . . . . . . . . 25
2.2 Out-of-Sample Forecasting, 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 . . . . 118
5.4 LIBOR Spreads . . . . . . . . . . . . . . . . . . . 122
List of Tables

1.1 Yield Statistics ........................................... 6
1.2 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 RMSFE, Four Models ................. 76
3.3 Out-of-Sample RMSFEs, Three Models ............... 78
To our wives
Comments Most Welcome
Preliminary Draft
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 (simple 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.

For helpful comments at various stages we thank – with-
out 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, Leo Krippner, Jose Lopez, Andre Lucas, Emanuel Mönch, James Morley, Charles Nelson, Ken Singleton, Dongho Song, Chuck Whiteman, and Tao Wu. 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 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 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@sf.frb.org
http://www.frbsf.org/economics/economists/grudebusch
Additional Acknowledgment

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


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 re-
lationships among them: the discount curve, the forward rate curve,
and the yield curve. Let \( P(\tau) \) denote the price of a \( \tau \)-period discount
bond, i.e., the present value of \$1 receivable \( \tau \) periods ahead. If \( y(\tau) \)
is its continuously compounded yield to maturity, then by definition
\[
P(\tau) = e^{-\tau y(\tau)}.
\]
Hence the discount curve and yield curve are immediately and fund-
amentally related. Knowledge of the discount function lets one
calculate the yield curve.

The discount curve and the forward rate curve are similarly fund-
amentally related. In particular, the forward rate curve is defined
as
\[
f(\tau) = \frac{-P'(\tau)}{P(\tau)}.
\]
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,
\[
y(\tau) = \frac{1}{\tau} \int_0^\tau f(u)du.
\]
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 zero-coupon yields at fixed standardized maturities (e.g., six-month, ten-year, ...). Hence we now provide some brief 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).

The second 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. It has complications of its own, however, as the implied forward rates are not necessarily positive.

An alternative and popular approach to yield construction is due to Fama and Bliss (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

\(^1\)See also McCulloch and Kwon (1993).
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 smooths the 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. 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. 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 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, mean yields (the "average yield curve") increase with

---

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

3 The 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 typically it does not address dynamics and, moreover, many special aspects of yield curve modeling lead us to rather different approaches.
1.3. YIELD CURVE FACTS

Figure 1.1: 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, from January 1985 through December 2008. Data are from the Board of Governors of the Federal Reserve System.
CHAPTER 1. FACTS, FACTORS, AND QUESTIONS

Table 1.1: Yield Statistics

<table>
<thead>
<tr>
<th>Maturity (Months)</th>
<th>( \bar{y} )</th>
<th>( \hat{\sigma}_y )</th>
<th>( \hat{\rho}_y(1) )</th>
<th>( \hat{\rho}_y(12) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>4.9</td>
<td>2.1</td>
<td>0.98</td>
<td>0.64</td>
</tr>
<tr>
<td>12</td>
<td>5.1</td>
<td>2.1</td>
<td>0.98</td>
<td>0.65</td>
</tr>
<tr>
<td>24</td>
<td>5.3</td>
<td>2.1</td>
<td>0.97</td>
<td>0.65</td>
</tr>
<tr>
<td>36</td>
<td>5.6</td>
<td>2.0</td>
<td>0.97</td>
<td>0.65</td>
</tr>
<tr>
<td>60</td>
<td>5.9</td>
<td>1.9</td>
<td>0.97</td>
<td>0.66</td>
</tr>
<tr>
<td>120</td>
<td>6.5</td>
<td>1.8</td>
<td>0.97</td>
<td>0.68</td>
</tr>
</tbody>
</table>

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.

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 persistent. As with yields, the one-month spread autocorrelations are very large, but they decay more quickly, so that the twelve-month spread autocorrelations are noticeably smaller than those for yields.
1.4. YIELD CURVE FACTORS

Table 1.2: Spread Statistics

<table>
<thead>
<tr>
<th>Maturity (Months)</th>
<th>$\bar{s}$</th>
<th>$\hat{s}_s$</th>
<th>$\hat{\rho}_s(1)$</th>
<th>$\hat{\rho}_s(12)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>-1.6</td>
<td>1.3</td>
<td>0.98</td>
<td>0.44</td>
</tr>
<tr>
<td>12</td>
<td>-1.4</td>
<td>1.1</td>
<td>0.98</td>
<td>0.46</td>
</tr>
<tr>
<td>24</td>
<td>-1.1</td>
<td>0.9</td>
<td>0.97</td>
<td>0.48</td>
</tr>
<tr>
<td>36</td>
<td>-0.9</td>
<td>0.7</td>
<td>0.97</td>
<td>0.47</td>
</tr>
<tr>
<td>60</td>
<td>-0.6</td>
<td>0.4</td>
<td>0.96</td>
<td>0.44</td>
</tr>
<tr>
<td>120</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

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. The sample period is January 1985 through December 2008.

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. Fortunately, it turns out that financial markets typically conform to a certain type of restricted vector autoregression, displaying factor structure.

Factor structure relates to 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.

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.
CHAPTER 1. FACTS, FACTORS, AND QUESTIONS


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, because it effectively imposes perfect correlation between unanticipated movements in yields at different maturities, 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—involves 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 three-factor representations emerge from arbitrage-free, dynamic term structure models...for a wide range of maturities.”

Typically three factors, or principle components, are all that one needs to explain most yield variation. In our data set the first three

\(^4\)Interestingly, 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).

\(^5\)For now we do not distinguish between government and corporate bond yields. We will consider credit risk spreads later.
1.4. **YIELD CURVE FACTORS**

Figure 1.2: Yields in Two Dimensions

![Yield vs Year Graph]

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. The sample period is January 1985 through December 2008.

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 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.
CHAPTER 1. FACTS, FACTORS, AND QUESTIONS

Figure 1.3: Yield Principal Components

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

Figure 1.4 shows that the three bond yield factors effectively capture level, slope, and curvature. We plot the three factors against standard empirical level, slope and curvature proxies (the 10-year yield, the 10Y-6M spread, and a 6M+10Y-2*5Y butterfly spread, respectively). This is important because different and quite specific features of the macroeconomy are likely relevant for the different factors, once we recognize that they are effectively level, slope, and curvature. Inflation, for example, is clearly related to the yield curve level, and the stage 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
1.4. YIELD CURVE FACTORS

Table 1.3: Yield Principal Components Statistics

<table>
<thead>
<tr>
<th>PC</th>
<th>$\sigma$</th>
<th>$\hat{\rho}(1)$</th>
<th>$\hat{\rho}(12)$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>First</td>
<td>2.35</td>
<td>0.97</td>
<td>0.67</td>
<td>0.98</td>
</tr>
<tr>
<td>Second</td>
<td>0.52</td>
<td>0.97</td>
<td>0.45</td>
<td>0.95</td>
</tr>
<tr>
<td>Third</td>
<td>0.10</td>
<td>0.83</td>
<td>0.15</td>
<td>0.70</td>
</tr>
</tbody>
</table>

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. The sample period is January 1985 through December 2008.

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 cancels out for 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 Dungey et al. (2000), Giese (2008) and Jardet et al. (2010), among others. Alternatively, 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.
Figure 1.4: Empirical Level, Slope, and Curvature, and First Three Principal Components

Notes: We show the standardized empirical level, slope, and curvature with dark lines, and the first three standardized principal components with lighter lines.
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 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 highly accurate empirical descriptions 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 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 didn’t have factor structure, but
again, they do! Hence we’re in a most fortunate situation. We need low-dimensional factor structure for statistical tractability, and mercifully, the data conform tightly.

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. 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.

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.

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

---

7See Diebold (2007) for additional discussion of the parsimony principle.
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 entire yield curve at any point in time, say, for pricing derivatives.
1.5. **YIELD CURVE QUESTIONS**

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 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 forms for the latent factors and derives the associated restrictions on factor loadings that ensure absence of arbitrage.

### 1.5.3 Is imposition of absence of 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 any credible yield curve model must display freedom from arbitrage. But our models are just models, and all models are false. Things change 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. That is, if the underlying factor
model is misspecified badly enough, no-arbitrage restrictions may actually degrade empirical performance.

Conversely, a model may admit arbitrage yet provide a good approximation to a much more complicated (perhaps arbitrage-free) 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. 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. The point is that we don’t force restrictions to hold exactly, but rather we nudge things in their direction. An obvious benchmark shrinkage direction is 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 DNS 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 can lead to improved predictive 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 investors who also demand a large equity premium as compensation for holding risky stocks. Furthermore, as suggested by many statistical tests in the literature, these risk premiums on nominal bonds appear to vary over time, contradicting the assumption of risk
1.5. **YIELD CURVE QUESTIONS**

neutrality (e.g., Campbell and Shiller (1991), Cochrane and Piazzesi (2005)).

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.” The large literature on stochastic volatility takes the former approach, allowing the variability of the 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, which translate a unit of factor volatility into a term premium.

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 with no economic interpretation. 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, 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

---

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

10 Some 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.
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 for exploration. However, it is striking how little interchange or overlap between the two research literatures has occurred in the past. Notably, both the Nelson-Siegel and affine no-arbitrage dynamic latent factor models provide useful statistical descriptions of the yield curve, but they offer little insight into the nature of the underlying economic forces that drive its movements. To shed some light on the fundamental determinants of interest rates, researchers have begun to incorporate macroeconomic variables into these yield curve models.

For example, Diebold et al. (2006b) provide a macroeconomic interpretation of the Nelson-Siegel 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 Nelson-Siegel 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.\textsuperscript{11} They find that output shocks have a significant impact on intermediate yields and curvature, while inflation surprises have large effects on the level of the entire yield curve.

\textsuperscript{11}To avoid relying on specific macro series, Ang and Piazzesi construct their measures of real activity and inflation as the first principal component of a large set of candidate macroeconomic series,
1.5. YIELD CURVE QUESTIONS

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.
CHAPTER 1. FACTS, FACTORS, AND QUESTIONS

Comments Most Welcome
Chapter 2

Dynamic Nelson-Siegel

Here we begin our journey. We start with static curve-fitting in the cross-section, but we proceed quickly to dynamic modeling with all its nuances and opportunities (and pitfalls).

2.1 Curve Fitting

As we will see, Nelson-Siegel fits a smooth yield curve to unsmoothed yields. One can arrive at a smooth yield curve in a different way, fitting a smooth discount curve to unsmoothed bond prices and then inferring the implied yield curve. That’s how things developed historically, but there are problems, as discussed in Chapter 1.

So let us proceed directly to the static Nelson-Siegel representation. At any time, one sees a large set of yields and may want to fit a smooth curve. Nelson and Siegel (1987) begin with a forward rate curve and fit the function

\[ f(\tau) = \beta_1 + \beta_2 e^{-\lambda \tau} + \beta_3 \lambda \tau e^{-\lambda \tau}. \]

The corresponding static Nelson-Siegel yield curve is

\[ y(\tau) = \beta_1 + \beta_2 \left( \frac{1 - e^{-\lambda \tau}}{\lambda \tau} \right) + \beta_3 \left( \frac{1 - e^{-\lambda \tau}}{\lambda \tau} - e^{-\lambda \tau} \right). \]  \hspace{1cm} (2.1)

Note well that these are simply functional form suggestions for fitting the cross section of yields.
Chapter 3

Arbitrage-Free
Nelson-Siegel

Because bonds trade in deep and well-organized markets, the theoretical restrictions that eliminate opportunities for riskless arbitrage across maturities and over time hold powerful appeal, and they provide the foundation for a large finance literature on arbitrage-free models that started with Vasicek (1977) and Cox et al. (1985). Those models specify the risk-neutral evolution of the underlying yield curve factors as well as the dynamics of risk premia. Following Duffie and Kan (1996), the affine versions of those models are particularly popular, because yields are convenient affine functions of underlying latent factors with factor loadings that can be calculated from a simple system of differential equations.

Unfortunately, the canonical affine arbitrage-free models often exhibit poor empirical time series performance, especially when forecasting future yields, a point forcefully made by Duffee (2002). In addition, and crucially, the estimation of those models is known to be problematic, in large part because of the existence of numerous likelihood maxima that have essentially identical fit to the data but very different implications for economic behavior. These empirical problems appear to reflect a pervasive model overparameterization,

\footnote{Piazzesi (2010) provides an insightful survey of affine arbitrage-free term structure models.}
and as a solution, many researchers (e.g., Duffee (2002) and Dai and Singleton (2002)) simply restrict to zero those parameters with small t-statistics in a first round of estimation. Joslin et al. (2010) put it well:

Faced with such a large number of free parameters, standard practice has been to estimate a maximally flexible dynamic term structure model, set to zero many of the parameters...that are statistically insignificant at a conventional significance level, and then to analyze the constrained model. [p. 14]

The resulting more parsimonious structure is typically somewhat easier to estimate and has fewer troublesome likelihood maxima. However, the additional restrictions on model structure are not well motivated theoretically or statistically, and their arbitrary application and the computational burden of estimation effectively preclude robust model validation and thorough simulation studies of the finite-sample properties of the estimators.

In this chapter, we discuss a new class of affine arbitrage-free models that overcome the problems with empirical implementation of the canonical affine arbitrage-free model. This new class is based on DNS and retains its empirical tractability. Thus, from one perspective, we take the theoretically rigorous but empirically problematic affine arbitrage-free model and make it empirically tractable by incorporating DNS elements.

From another perspective, we take the DNS model and make it theoretically more satisfactory. Let us elaborate. As we have emphasized, DNS is simple and stable to estimate, and it is quite flexible and fits both the cross section and time series of yields remarkably well. Theoretically, DNS imposes certain economically desirable properties, such as requiring the discount function to approach zero with maturity, and as we have shown, it corresponds to a modern three-factor model of time-varying level, slope, and curvature. However, despite its good empirical performance and a certain amount of theoretical appeal, DNS fails on an important theoretical dimension: it does not impose the restrictions necessary to eliminate
Chapter 4

Extensions

In this chapter, we highlight aspects of the vibrant ongoing research program associated with the ideas developed in earlier chapters. We begin with a collage-style sketch of work involving Bayesian analysis, functional form for factor loadings, term structures of credit spreads, and nonlinearitys. We then discuss in greater detail a time-honored topic that has received attention both historically and presently, incorporation of more than three yield factors. Third, we treat stochastic volatility in both DNS and AFNS environments, with some attention to the issue of unspanned stochastic volatility. Finally, we also discuss in detail a crucially important topic for the emerging research agenda, the incorporation of macroeconomic fundamentals in their relation to bond yields. Related, we introduce aspects of modeling real vs. nominal yields in DNS/AFNS environments, a theme that we treat in detail in Chapter 5.

4.1 Variations on the Basic Theme

DNS/AFNS research is moving forward at a steady pace. Here we sketch a few of the more interesting and important developments, to convey a feel for the issues and ideas, and to provide relevant references for those who want to dig deeper.
4.1.1 Bayesian shrinkage

Thus far we have either imposed various constraints, or not. Examples include the independence constraints associated with independent-factor DNS, the no-arbitrage constraints associated with correlated-factor AFNS, or the simultaneous imposition of both as in independent-factor AFNS.

One can think of such “hard constraints” as the outcome of a Bayesian analysis with spiked priors. But one may want to impose “soft constraints,” coaxing (“shrinking”) estimates in certain directions without forcing them. That is, one may want to do an informative-prior Bayesian analysis, but with less-than-perfect prior precision, so that likelihood information is blended with prior information rather than simply discarded. As is well-known, such shrinkage often improves forecasts.¹

There is ample opportunity to use shrinkage ideas in predictive yield curve modeling, and research along those lines is appearing. Carriero et al. (2010) perform a Bayesian analysis of a vector autoregression for a large set of yields, using the “Minnesota prior” popular in the empirical macroeconomics literature.² They obtain encouraging forecasting results, systematically beating random-walk forecasts, although not by much.

The Minnesota prior is a statistical prior. The obvious benchmark economic shrinkage direction for a yield curve model, however, is toward the restrictions implied by absence of arbitrage. Carriero (2011) does just that, shrinking but not forcing a vector autoregression toward the no-arbitrage configuration of Ang and Piazzesi (2003), with striking results. Forecasts based on vector autoregressions shrunk toward Ang-Piazzesi no-arbitrage improve on unrestricted vector-autoregressive forecasts, and on Ang-Piazzesi restricted vector-autoregressive forecasts, and on vector-autoregressive forecasts shrunk in other directions. In particular, forecasts from Ang-Piazzesi shrink vector autoregressions dominate those from vector-random-walk shrunk vector autoregressions associated with

---

¹See, for example, Diebold (2007).
²For discussion of the Minnesota prior, see the classic Doan et al. (1984).
Chapter 5

Macro-Finance

From the vantage point of incorporating macroeconomic considerations into yield curve modeling, one can view the approaches introduced previously in section 4.4 as preparatory, paving the way for more extensive explorations. In this chapter, we move in that direction, discussing a variety of AFNS macro-finance yield curve approaches.

5.1 Macro-Finance Yield Curve Modeling

A key feature of the global financial crisis and recession of 2007-2009 was the close feedback between the real economy and financial conditions. In many countries, the credit and asset price boom that preceded the crisis went hand in hand with strong spending and production. Similarly, during the crash, deteriorating financial conditions both contributed to the deep declines in economic activity and were exacerbated by them. Still, modeling this close feedback poses a significant challenge to macroeconomists and finance economists because of the long-standing separation between the two disciplines. In macro models, the entire financial sector is often represented by a single interest rate with no accounting for credit or liquidity risk and no role for financial intermediation or financial frictions. Similarly,
finance models often focus on the consistency of asset prices across markets with little regard for underlying macroeconomic fundamentals. To understand important aspects of the recent financial crisis and, more generally, the intertwined dynamics of interest rates and the economy, a joint macro-finance perspective is likely necessary.

Of course, differences between the finance and macro perspectives reflect in part different questions of interest and different avenues for exploration; however, it is striking that there is so little interchange or overlap between the two research literatures. At the very least, it suggests that there may be synergies from combining elements of each. From a finance perspective, the short rate is a fundamental building block for rates of other maturities, because long yields are driven in significant part by expected future short rates. From a macro perspective, the short rate is a key monetary policy instrument, which is adjusted by the central bank to achieve economic stabilization goals. Taken together, a joint macro-finance perspective would suggest that understanding the central bank policy response to fundamental macroeconomic shocks should explain movements in the short end of the yield curve; furthermore, with the consistency between long and short rates enforced by the no-arbitrage assumption, expected future macroeconomic variation should account for movements farther out in the yield curve as well.

One key strand of macro-finance research examines the finance implications for bond pricing in a model with macroeconomic variables. As a theoretical matter, the term premium on a long-term nominal bond compensates investors for inflation and consumption risks over the lifetime of the bond. A large finance literature finds that these risk premiums are substantial and vary significantly over time (e.g., Campbell and Shiller (1991), Cochrane and Piazzesi (2005)). However, Backus et al. (1989) find the standard consumption-based asset pricing model of an endowment economy cannot account for such large and variable term premiums. The basic inability of a

---

1Even a default-free nominal bond is risky if its price covaries with the bondholder’s marginal utility of consumption. If times of high inflation are correlated with times of low output, then a nominal bond loses value just when the investor values consumption the most, so it would carry a risk premium.
Chapter 6

Epilogue

We begin with a bit of history, as what we have done in this book blends the old and the new.

Early on, bond markets were few and far between, so the term structure of bond yields remained a dormant issue. But as financial markets developed, and as economic theory and measurement advanced, issues related to the term structure emerged as central in asset allocation, asset pricing and risk management.

The earliest work effectively assumes perfect foresight, as in Bohm-Bawerk (1889). Building on that work, early interwar term structure modeling typically allows for risk but assumes risk neutrality, as in Fisher (1930) and Keynes (1936), culminating in the classic expectations theory of Hicks (1946). Hicks’ most basic theory, sometimes called the “traditional expectations theory,” asserts that current forward rates equal expected future spot rates, in which case current long rates are simple averages of expected future spot rates.

An immediate implication of the traditional expectations theory is that the yield curve will be flat if spot rates are expected to remain unchanged.\(^1\) This conflicts with the observed upward-sloping average yield curve and directs attention to the possibility of “term premia” that may separate forward rates from expected future spot rates.

\(^1\)This would occur, for example, if the spot rate is a martingale and expectations are formed rationally, in which case the optimal spot rate forecast at all horizons is “no change.”
The first resulting variant of the traditional expectations theory is the “liquidity preference theory” of Hicks (1946), which posits that lenders prefer short maturities (to avoid the risk associated with holding long-duration bonds) and borrowers prefer long maturities (to lock in the cost of finance), so that lenders require and receive long-maturity premia from borrowers. That is, the yield curve slopes upward on average.

Richer approximations to average yield curves, and time-varying conditional curves, soon followed. In the “preferred habitat” or “market segmentation” theory of Modigliani and Sutch (1967), different agents prefer to borrow or lend in different regions of the curve. Depending on the distribution of agents among various preferred habitats, the curve can take many shapes, not just upward sloping. Moreover, the “slope factor” generated by preferred habitats can vary over time, as the distribution of agents among various preferred habitats can vary over time, for example with the level of yields and/or with expected business conditions, as in Meiselman (1962), Kessel (1965), Van Horne (1965), and Nelson (1972). Perhaps most notably in that tradition, Nelson (1972) identified expected spot rates via time series methods, subtracted them from current forward rates to get term premia, and then projected the term premia on business conditions indicators.

The “traditional” literature’s concern with dynamic yield curve evolution is very much maintained in the “modern” literature, which in many respects began with Vasicek (1977), and which was significantly extended by classic subsequent contributions by Duffie and Kan (1996), Dai and Singleton (2000), Ang and Piazzesi (2003), among others. The key distinguishing characteristic of the modern literature is explicit enforcement of absence of arbitrage.

Interestingly, DNS and AFNS in many respects bridge the tra-

---

2 In addition to theory, measurement also advanced rapidly during this time. High-quality yield curve data, for example, was constructed by Durand (1958) and Malkiel (1966), among others.

3 Note that the level of yields is of course one aspect of business conditions, so the key issue is relating the dynamics of the term structure to the dynamics of the business cycle. The early literature’s emphasis on linking term premia to yield levels evidently traces to Keynes (1936), pp. 201-202).
Bibliography


BIBLIOGRAPHY


Bohm-Bawerk, E. (1889), *Capital and Interest*, Macmillan.


BIBLIOGRAPHY


BIBLIOGRAPHY


BIBLIOGRAPHY


University of Amsterdam, Tinbergen Institute, and Erasmus University Rotterdam.


BIBLIOGRAPHY


*NBER Macro Annual*, 389–442.


Williams, David (1997), Probability with Martingales, Cambridge University Press.
BIBLIOGRAPHY


