Information in (and not in) the term structure

** WARNING. Paper is highly preliminary. **

Gregory R. Duffee∗
Haas School of Business
University of California – Berkeley
First draft: March 2008

ABSTRACT

Casual intuition says that today’s term structure reflects all information investors have about expected future yields. However, this is not required by finance theory, nor is it consistent with observed Treasury yield behavior. Kalman filter estimation uncovers a factor that has an almost imperceptible effect on yields, but has clear forecast power for future short-term interest rates and substantial forecast power for future excess bond returns. The factor appears to be related to short-run fluctuations in economic activity.

∗Voice 510-642-1435, email duffee@haas.berkeley.edu. Address correspondence to 545 Student Services Building #1900, Berkeley, CA 94720.
1 Introduction

Investors use all relevant available information to price bonds. Since bonds’ cash flows are fixed, prices are determined by investors’ required expected returns, and thus by investors’ expectations of future bond prices. This logic appears to imply that today’s term structure contains all information relevant to forecasting future yields.

Researchers commonly invoke this logic when building and estimating multifactor term structure models. For example, properties of the unconditional covariance matrix of bond yields are often used to justify the number of factors assumed to drive term structure dynamics. Models are frequently estimated by assuming a one-to-one mapping from the factors to an equal number of bond yields. Both choices implicitly rely on the notion that the cross-section of bond yields follows a Markov process. We know the assumption is not literally true, because of strong evidence that yields on individual bonds have small idiosyncratic components associated with market imperfections. The core idea is that major determinants of expected future yields should also be major determinants of current yields.

Two empirical observations by Cochrane and Piazzesi (2005) cast some doubt on this view. First, they find that the forward rate from year four to year five contains substantial information about future excess bond returns, even though the contribution of this forward rate to the overall volatility of cross-section of bond yields is very small. Second, they find that lagged bond yields contain information about future excess bond returns not found in current bond yields. The noise in individual bond yields presumably plays a role in these results, as suggested by Cochrane and Piazzesi. Nonetheless, we need a reasonable theoretical interpretation of this wedge between determinants of the cross-section and determinants of expectations.

This paper shows that from a formal perspective, it is easy to build a multifactor model in which one of the factors plays an important role in determining investors’ expectations of future yields, yet has zero effect on current yields. The factor must have opposite effects on expected future interest rates and bond risk premia. Consider, for example, economic news that raises risk premia and simultaneously leads investors to believe the Fed will soon cut short-term interest rates. The increase in risk premia induces an immediate increase in long-term bond yields, while the expected drop in short rates induces an immediate decrease in these yields. In a Gaussian term structure model, a single parameter restriction equates these effects, leaving the current term structure—but not expected future term structures—unaffected by the news. More generally, factors that drive risk premia and expected short rates in opposite directions will have smaller effects on the cross-section of yields than they will have on yield dynamics.
This conclusion has implications for building and estimate term structure models. The choice of number of factors in a model should not be based on the number of factors that appear to explain the cross-section of yields. In addition, models should be estimated using techniques that do not rely on the ability to infer time- \( t \) factors from time- \( t \) yields. Even if the factors do not literally have a zero effect on the time- \( t \) term structure, their effect on yields may be too small to readily distinguish from idiosyncratic components. Based on this reasoning, I fit a five-factor Gaussian term structure model to monthly Treasury yields over the period 1964 through 2006. Estimation is with the Kalman filter, which allows us to infer the presence of factors from both dynamics and the cross section.

Estimation uncovers a term structure factor that has a trivial effect on the cross-section of Treasury yields but contains substantial information about both expected future short rates and—necessarily—expected excess bond returns. Based on the model’s point estimates, a one standard deviation change in the factor has an almost imperceptible effect on the term structure (on the order of a few basis points), lowers the expected one-year-ahead short rate by about 35 basis points, and raises the expected excess return to a five-year bond over the next year by about 1.4 percent. This “expectation” factor accounts for about 30 percent of the total variation in expected excess bond returns.

A skeptic could argue that this estimation uncovers a factor ex-post. In other words, the reason the factor does show up in the term structure is because investors at the time did not observe it. However, evidence from the Survey of Professional Forecasters confirms that survey-based expectations of future short rates move contemporaneously with filtered estimates of the factor. The factor appears to be related to short-run fluctuations in economic activity. An increase in the factor corresponds to lower expected future short rates, higher risk premia, and lower expected growth in industrial production.

I also investigate properties of regressions that use the term structure to forecast future excess annual bond returns. Under the maintained hypothesis that the estimated five-factor model is correct, such regressions are incapable of capturing all of the true variation in expected excess returns because they cannot capture fully the expectations factor. The Cochrane and Piazzesi (2005) regression that uses five forward rates as forecasting variables slightly outperforms a regression that uses measures of level, slope, and curvature. In population, the difference in \( R^2 \) is about one percentage point. It also slightly outperforms in long (43-year) finite samples. However, in these finite samples it is also easy to conclude, mistakenly, that the two regressions capture substantially different amounts of information. Although the mean difference in these finite-sample \( R^2 \)’s is only one percentage point, there is a twelve percentage point range from one end of the 95th percentile bound to the other.

The term structure model is presented in the next section. Section 3 summarizes prop-
erties of the estimated model. Section 4 compares the expectation factor to survey evidence on expectations and links the factor to the macroeconomy. Finite-sample properties on forecasting regressions are in Section 5. Concluding comments are in Section 6.

2 The modeling framework

The objective of this section is to explain why the important determinants of the cross-section of bond yields need not correspond to the important determinants of yield dynamics. To make this point in the starkest terms, I build a model in which \( n \) factors are necessary to model term structure dynamics, but only \( n - 1 \) factors appear in the term structure.

I follow much of the modern term structure literature by abstracting from standard economic concepts such as utility functions and consumption dynamics. Instead, both the short rate and the nominal pricing kernel are functions of a latent state vector. The factors and their dynamics can be viewed as reduced-form representations of inflation, business cycles, and market clearing.

2.1 The standard Gaussian model

I use a standard discrete time Gaussian term structure framework. The use of discrete time is innocuous. The role played by the Gaussian assumption is discussed in Section 2.5. The one-period interest rate is \( r_t \). This rate is continuously compounded and expressed per period. (For example, if a period is a month, \( r_t = 0.01 \) corresponds to twelve percent/year.) Interest rate dynamics are driven by a length-\( n \) state vector \( x_t \). The relation between the short rate and the state vector is

\[
r_t = \delta_0 + \delta_1 x_t. \tag{1}
\]

The state vector has first-order Markov dynamics

\[
x_{t+1} = \mu + K x_t + \Sigma \epsilon_{t+1}, \quad E_t (\epsilon_{t+1} \epsilon_{t+1}') \sim N(0, I). \tag{2}
\]

The period-\( t \) price of a zero-coupon bond that pays a dollar at \( t + m \) is denoted \( P^{(m)}_t \). The corresponding continuously-compounded yield is \( y^{(m)}_t \). Bond prices satisfy the law of one price

\[
P^{(m)}_t = E_t \left( M_{t+1} P^{(m-1)}_{t+1} \right) \tag{3}
\]

where \( M_{t+1} \) is the pricing kernel. The pricing kernel has the log linear form

\[
\log M_{t+1} = -r_t - \Lambda' \epsilon_{t+1} - \frac{1}{2} \Lambda' \Lambda t. \tag{4}
\]
The vector $\Lambda_t$ is the compensation investors require to face shocks to state vector. The price of risk satisfies

$$\Sigma \Lambda_t = \lambda_0 + \Lambda_1 x_t,$$

which is the essentially affine form introduced in Duffee (2002). Bonds are priced using the equivalent-martingale dynamics

$$x_{t+1} = \mu^q + K^q x_t + \Sigma \epsilon_t^q,$$

where the equivalent-martingale parameters are

$$\mu^q = \mu - \lambda_0, \quad K^q = K - \lambda_1.$$

The discrete-time analogues of the restrictions in Duffie and Kan (1996) imply that zero-coupon bond yields can be written as

$$y_t^{(m)} = A_m + B'_m x_t,$$

where the scalar $A_m$ and the $n$-vector $B_m$ are functions of the parameters in (1) and (6). The focus of this paper is on yield factor loadings, which can be written as

$$B'_m = \frac{1}{m} \delta'_1 (I + K^q + (K^q)^2 + \cdots + (K^q)^{m-1})$$

$$= \frac{1}{m} \delta'_1 (I - K^q)^{-1} (I - (K^q)^m).$$

### 2.2 The information in the term structure

In the absence of specific parameter restrictions, the period-$t$ state vector can be inferred from a cross-section of period-$t$ bond yields. Stack the yields on $n$ zero-coupon bonds in the vector $y^a_t$. We can write this vector as

$$y^a_t = A^a + B^a x_t$$

where $A^a$ is a length-$n$ vector containing $A_m$ for each of the $n$ bonds and $B^a$ is a square matrix with rows $B'_m$ for each bond. In general, $B^a$ is invertible. Put differently, element $i$ of the state vector affects the $n$ bond yields in a way that cannot be duplicated by a combination of the other elements. With invertibility, the term structure contains the same
information as \( x_t \). We can write

\[
x_t = (B^a)^{-1} (y_t^a - A^a).
\] (11)

Since \( x_t \) is Markov and the term structure of yields contains the same information as \( x_t \), the term structure is also first-order Markov.

We now investigate special cases of this Gaussian framework where \( B^a \) has rank less than \( n \), so that the state vector cannot be extracted from the term structure. An example illustrates the mathematics and the economic intuition.

### 2.3 A two-factor example

Consider the two-factor Gaussian model. Because the latent factors in this model can be arbitrarily rotated, the state vector can be transformed into the short rate and some other factor, denoted \( f_t \). For this rotation, the dynamics of the state vector are (explicitly indicating the elements of the feedback matrix)

\[
\begin{pmatrix}
  r_{t+1} \\
  f_{t+1}
\end{pmatrix}
= \mu
+ \begin{pmatrix}
  k_{11} & k_{12} \\
  k_{21} & k_{22}
\end{pmatrix}
\begin{pmatrix}
  r_t \\
  f_t
\end{pmatrix}
+ \Sigma \epsilon_{t+1}.
\] (12)

When \( k_{12} \) is not restricted to zero, time-\( t \) expectations of future short rates depend on both \( r_t \) and \( f_t \). Thus we can think of \( f_t \) as all information about future short rates that is not captured by the current short rate.

If investors were risk-neutral, the level of \( f_t \) would necessarily affect the term structure through expectations of future changes in the short rate. But if risk premia also vary with \( f_t \), the net effect of \( f_t \) on yields is ambiguous. The restriction adopted in this example is that changes in risk premia exactly cancel expectations of future short rates, leaving yields unaffected by \( f_t \). Formally, the requirement is \( k_{12}^q = 0 \), or \( k_{12} = \lambda_{1(12)} \). Then the equivalent-martingale dynamics of the state are

\[
\begin{pmatrix}
  r_{t+1} \\
  f_{t+1}
\end{pmatrix}
= \mu^q
+ \begin{pmatrix}
  k_{11}^q & 0 \\
  k_{21}^q & k_{22}^q
\end{pmatrix}
\begin{pmatrix}
  r_t \\
  f_t
\end{pmatrix}
+ \Sigma^q \epsilon_{t+1}.
\] (13)

A glance at (13) reveals that under the equivalent-martingale measure, the short rate follows a (scalar) first-order Markov process. The loading of the \( m \)-period bond yield on the state vector is, from (9),

\[
B_m = \begin{pmatrix}
  \frac{1}{m} (1 - k_{11}^q)^{-1} (1 - (k_{11}^q)^m) \\
  0
\end{pmatrix}.
\] (14)
Thus the matrix $B^a$ in (10) cannot be inverted because it has a column of zeros. The factor $f_t$ cannot be inferred from the period-$t$ term structure.

Although the factor does not affect yields, investors observe it. They take it into account when setting bond prices and forming expectations of future yields (or equivalently, future returns to holding bonds). For concreteness, consider the case $k_{12} > 0$. Then for fixed $r_t$, an increase in $f_t$ raises investors’ expectations of future short rates. For example, consider macroeconomic news, such as unexpectedly high GDP growth, that raises the likelihood of future tightening by the Federal Reserve. If investors’ willingness to bear interest risk did not change with $f_t$, this news would raise current long-maturity bond yields. But with the restriction $k_{12} = \lambda_{1(12)}$, investors accept lower expected excess bond returns. The change in willingness to bear risk offsets exactly the news about expected future short rates, leaving yields unaffected.

The functional relation between expected excess returns and $f_t$ can be seen in the formula for the expected excess log return, from $t$ to $t + 1$, on a bond with maturity $m$ at period $t$. (Here, “excess” is in excess of the short rate.) The period-$t$ expectation is

$$E_t \left( x r^{(m)}_{t,t+1} \right) \equiv m y^{(m)}_t - (m - 1) E_t \left( y^{(m-1)}_{t+1} \right) - r_t$$

$$= m A_m - (m - 1) A_{m-1}$$

$$+ (1 - k^q_{11})^{-1} \left[ (1 - (k^q_{11})^m) - \left( 1 - (k^q_{11})^{(m-1)} \right) k_{11} - 1 \right] r_t$$

$$- (1 - k^q_{11})^{-1} \left( 1 - (k^q_{11})^{(m-1)} \right) k_{12} f_t.$$ (15)

The final term in (15) captures the dependence of expected excess returns on $f_t$.

Even if an econometrician knows the parameters of the model, she cannot infer $f_t$ from the cross-section of yields at $t$. Nor can $f_t$ be backed out of the price of some other fixed-income instrument, such as bond options. The econometrician can, however, use a panel of data to form filtered estimates of $f_t$. The filtering approach is discussed again in Section 3.2. The intuition behind filtering is easier to grasp if we call it learning by the econometrician. The period-$t$ forecast error (the difference between realized yields and the econometrician’s $t-1$ forecast) is produced by both true period-$t$ shocks and the error in the econometrician’s $t-1$ prediction of $f_{t-1}$. The cross-sectional pattern of the period-$t$ forecast errors helps the econometrician revise her prediction of $f_{t-1}$ and form her prediction of $f_t$.

In this example, the short rate follows a two-factor Markov process under the physical measure and a one-factor Markov process under the equivalent martingale measure. A single parameter restriction is required to generate this structure. Armed with the intuition of this example, it is straightforward to proceed to the more general case in which the short rate follows an $n$-factor Markov process under the physical measure and an $(n - 1)$-factor
Markov process under the equivalent martingale measure. As in the two-factor case, a single parameter restriction is required.

2.4 The $n$-factor version

Latent state vectors in affine term structure models are inherently arbitrary. Dai and Singleton (2000) describe in detail how they can be translated and rotated without observable consequences. One particular rotation simplifies considerably the analysis here. Beginning with the standard $n$-factor Gaussian model of Section 2.1, diagonalize the equivalent-martingale feedback matrix $K^q$ into

$$K^q = PV P^{-1}$$

where the columns of $P$ are eigenvectors and $V$ is a diagonal matrix of eigenvalues. Define a rotated state vector

$$x_t^* = Px_t.$$

The equivalent-martingale dynamics of the rotated state vector are

$$x_{t+1}^* = P\mu^q + Vx_t^* + P \Sigma \epsilon_{t+1}^q.$$  

With this rotation, each individual factor follows its own univariate first-order Markov process because $V$ is diagonal. Innovations among the factors can be correlated. The loading of the short rate on the rotated state vector is

$$(\delta_1^*)' = \delta_1' P^{-1}$$

Here, as in the two-factor case, a single parameter restriction produces a model where physical dynamics of the short rate follow an $n$-factor process and equivalent-martingale dynamics follow an $(n-1)$-factor process. The restriction is that for some $i$,

$$\delta_{1,i}^* = 0.$$  

This restriction implies that element $i$ of the state vector drops out of the equivalent-martingale dynamics of the short rate. It is immediate from (20) that the period-$t$ values of other $n-1$ factors are sufficient to determine the period-$t$ short rate. Similarly, the short rate at $t + \tau$ depends only on the period-$(t + \tau)$ values of $n-1$ factors. Since each factor follows a univariate Markov process under the equivalent-martingale measure, the period-$t$ equivalent-martingale expectation of the short rate at $t + \tau$ depends only on the period-$t$ values of those same $n-1$ factors. Therefore period-$t$ yields depend only $n-1$ factors.
As in the two-factor case, physical dynamics of the short rate depend on all \( n \) factors. The physical dynamics of the rotated state vector are

\[
x_{t+1}^* = P\mu + PKP^{-1}x_t^* + P\Sigma \epsilon_{t+1}.
\]  

(21)

As long as risk premia vary with the state vector \( (\lambda_1 \neq 0) \), the matrix \( P \) that diagonalizes \( K^n \) will not diagonalize \( K \). Then in general, each factor in the state vector contains information about the evolution of the short rate.

2.5 The role of the Gaussian setting

Section 2.4 shows that with an appropriate restriction on a term structure model, only \( n - 1 \) factors of an \( n \)-dimensional state vector affect bond yields. Models that exhibit unspanned stochastic volatility (USV), as described in Collin-Dufresne and Goldstein (2002), can be described similarly. Here I clarify the relation between the approach here and the USV approach.

Here, short rate dynamics are described by an \( n \)-factor Markov process under the physical measure and an \((n - 1)\)-factor Markov process under the equivalent martingale measure. All \( n - 1 \) factors that appear in the equivalent-martingale process affect bond yields. Thus we can say that under the equivalent-martingale measure, the term structure follows an \((n - 1)\) factor Markov process. By contrast, the USV framework is concerned only with the equivalent martingale measure. The physical measure is not specified. Under the equivalent martingale measure, the short rate follows an \( n \)-factor Markov process. Bond yields nonetheless do not depend on all \( n \) factors. (Prices of some other fixed-income instruments will depend on all \( n \) factors.) Thus under the equivalent-martingale measure, the term structure does not follow a Markov process.

The economic interpretations of the two sets of relevant parameter restrictions differ substantially. Here, variations in expected future short rates are offset by variations in risk premia. With USV, variations in equivalent-martingale expectations of future short rates are offset by variations in the Jensen’s inequality component of bond yields. Stochastic volatility is thus critical to USV models (hence the name of the model class), but does not appear here.

Although USV models appear to have little in common with the model here, they can provide an alternative mechanism driving a wedge between the factors driving dynamics of yields and the cross-section of yields. Set risk premia to zero so that physical and equivalent-martingale measures coincide. Then \( n \) factors are necessary to capture yield dynamics, while \( n - 1 \) factors affect bond yields. I do not pursue this approach because the parameter
restrictions necessary in a USV model are very tight.

In fact, one reason I use the Gaussian framework is to avoid complications associated with stochastic volatility. Reconsider the two-factor example of Section 2.3. If the conditional covariance matrix of factor innovations is allowed to be linear in \( f_t \) (a discrete-time approximation to a square-root diffusion model), then the level of \( f_t \) affects bond yields even when \( k_{12}^q = 0 \). Variations in risk premia can offset variations in expected future short rates, but do not offset variations in the Jensen’s inequality component of yields. This problem does not arise in the two-factor example if conditional variances are allowed to depend on the short rate instead of \( f_t \).

### 2.6 From theory to practice

Measurement error must be included in this model. Standard \( n \)-factor affine term structure models imply that zero-coupon bond yields are affine functions of an \( n \)-dimensional state vector. (Some of the coefficients may be zero.) A corollary is that the unconditional covariance matrix of \( d \) bond yields is singular for \( d > n \). Yet in the data, sample covariance matrices of zero-coupon bond Treasury yields are nonsingular for even large \( d \) (say, greater than ten). One interpretation of this result is that \( n \) is large, perhaps even infinite, as in Collin-Dufresne and Goldstein (2003). But from a variety of perspectives, it is more appealing to view bond yields as contaminated by small, transitory, idiosyncratic components. These components have three sources. First, there are market imperfections that distort bond prices, such as bid/ask spreads. Second, there are market imperfections that distort payoffs to bonds (and thus distort what investors will pay for bonds), such as special RP rates. Third, there are distortions created by the mechanical interpolation of zero-coupon bond prices from coupon bond prices.

Formally, a vector of \( d \) period-\( t \) yields on bonds with maturities \( m_1, \ldots, m_d \) is expressed as

\[
y_t = A + Bx_t + \eta_t, \quad \eta_t \sim \sigma^2_\eta N(0, I).
\]

For simplicity, in (22) the idiosyncratic components for each yield have the same variance. Element \( i \) of the vector \( A \) contains \( A_{m_i} \) and row \( i \) of the matrix \( B \) contains \( B'_{m_i} \).

The presence of measurement error broadens the relevance of the intuition behind the model here. It is probably unreasonable to assume that there is some factor for which variations in expected future short rates are exactly offset by variations in required expected returns. Although mechanically simple to construct in a no-arbitrage setting, the restriction is unlikely to be a natural consequence of a sensible equilibrium model. Thus there seems to be no theoretical justification to either \textit{a priori} impose the constraint or to test statistically
whether it is consistent with the data.

A more sensible interpretation of the model is that it explains why we should not look exclusively to the cross-section of bond yields when drawing inferences about term structure dynamics. It is easy to imagine types of news that have opposite effects on expected future short rates and investors’ required expected excess returns. For example, the Taylor (1993) rule and its variants (see, e.g., Clarida, Galí, and Gertler (2000)) suggest that good news about future output is also news that future short rates are likely to rise. If willingness to bear interest rate risk covaries positively with the business cycle, the immediate effect of such news on bond yields will not accurately reflect its importance in forecasting future short rates.

There are two practical lessons to take from this interpretation. First, there is no a priori reason to assume that all information relevant for forecasting yields or returns can be extracted from the cross-section. There may be some factor that is an important determinant of expected future yields that has only a tiny effect on the current term structure—so small that it is difficult or impossible to distinguish from the idiosyncratic components \( \eta_t \) in (22). Such a factor is better inferred using filtering techniques. A corollary is that formal term structure models written and estimated in terms of latent factors are preferable to models in which the factors are linear combinations of yields.

Second, when building formal term structure models, the choice of number of factors should not be made based on the number of factors that appear to explain the cross-section of yields. Instead, higher-dimension state vectors should be used to explore the possibility that some factor(s) are important for term structure modeling but have little effect on the cross-section of yields. The empirical work that follows explores these ideas in depth.

## 3 Empirical analysis

The core of this empirical analysis is the interpretation of an estimated five-factor Gaussian term structure model. The model is used in two ways. First, I ask whether there are factors that have little effect on the cross-section of yields yet are important for modeling dynamics. Second, assuming that the estimated model is correct, I study properties of regressions that forecast excess bond returns using the cross-section of bond yields.

### 3.1 Data

Treasury Bond yields are from the Center for Research in Security Prices (CRSP). The yield on a three-month Treasury bill is from the Riskfree Rate file (bid/ask average). Artificially-
constructed yields on zero-coupon bonds with maturities of one, two, three, four, and five years are from the Fama-Bliss file. Yields are observed at the end of each month from January 1964 through December 2006. The first observation is chosen to align with the sample studied by Cochrane and Piazzesi (2005).

Panel A of Table 1 reports means and standard deviations of the yields. Panel B displays the first five principal components of the unconditional covariance matrix of their yields. The principal components have the usual properties. The first principal component, which is roughly a level effect in the term structure, explains more than 97 percent of the total variance of the yields. The second and third principal components are slope and curvature effects respectively, and account for slightly more than 2 percent (slope) and less than 0.2 percent (curvature) of the total variance. The remaining principal components (including the unreported sixth component) account for only 0.04 percent of the total variance of yields.

Similar results popularized by Litterman and Scheinkman (1991) are typically used to motivate the number of factors included in formal term structure models. For example, the choice of three factors in Duffee (2002) is explicitly justified by this result. A major empirical question studied here is whether the relative importance of factors in the cross-section corresponds to their relative importance in dynamics.

3.2 Model estimation

A five-factor version of the Gaussian term structure model is estimated with maximum likelihood (ML). Because of the Gaussian structure, the Kalman filter produces correct conditional means and covariance matrices. The model of Section 2 is written in terms of dynamics under physical and equivalent-martingale measures. That form allows us to understand the economics underlying factor models of the term structure. However, for the purpose of estimation, it is convenient to use a slightly different parameterization.

Following the language of the Kalman filter, write the model in the form of a transition equation and a measurement equation. The state vector used in the estimated model is denoted $x_{t}^{\dagger}$. The transition equation is

$$x_{t+1}^{\dagger} = D^{\dagger} x_{t}^{\dagger} + \Sigma^{\dagger} \epsilon_{t+1}.$$  \hspace{1cm} (23)

In (23), $D^{\dagger}$ is a diagonal matrix and $\Sigma^{\dagger}$ is lower triangular with ones along the diagonal. The forms of $D^{\dagger}$ and $\Sigma^{\dagger}$ are normalizations, as is the state vector’s unconditional mean of zero. There are five latent factors in the state vector $x_{t}^{\dagger}$. Therefore there are a total of 15
free parameters in (23). The measurement equation is

\[ y_t = A + B^\dagger x_t^\dagger + \eta_t, \quad \eta_t \sim N(0, \sigma^2). \tag{24} \]

In (24), \( A \) is a \( 6 \times 1 \) vector and \( B^\dagger \) is a \( 6 \times 5 \) matrix. There is also a single standard deviation of measurement error, resulting in a total of 52 parameters.

Lurking behind the parameters of the measurement equation are the equivalent-martingale dynamics of \( x_t \). Because there are five factors to explain six bond yields, \( A \) and \( B^\dagger \) exactly identify the unconstrained parameters of the no-arbitrage model \( \delta_0, \delta_1, \mu^q, \) and \( K^q \). As discussed in Duffee (2008), numerical optimization of the likelihood function is faster and more reliable when the estimated parameters are \( A \) and \( B^\dagger \) than when they are the parameters of the no-arbitrage model. Here I follow exactly the optimization procedure used in that paper.

### 3.3 A principal components factor rotation

The state-vector rotation implied by (23) and (24) is convenient for estimation. A rotation based on principal components is more useful for interpreting the results. Denote ‘uncontaminated’ yields—yields without measurement error—by \( \tilde{y}_t \). Drop the three-year bond, denoting the vector of the remaining five yields by \( \tilde{y}_{\setminus 3,t} \). The loadings of these yields on the factors are denoted \( \tilde{B}_{\setminus 3}^\dagger \), a \( 5 \times 5 \) matrix. Estimates of the parameters of (23) and (24) imply a population covariance matrix of \( \tilde{y}_{\setminus 3,t} \). (As the model in Section 2.4 illustrates, there are parameterizations for which this covariance matrix is singular, but the parameter estimates do not happen to satisfy the restriction necessary for singularity.) Diagonalize this covariance matrix into

\[ \text{Var} (\tilde{y}_{\setminus 3,t}) = C_0 \Omega C_0^{-1}. \tag{25} \]

Define the \( 5 \times 5 \) matrix \( \Gamma \) as

\[ \Gamma = C_0^{-1} \tilde{B}_{\setminus 3}^\dagger. \tag{26} \]

The state vector that is easy to interpret is

\[ x_t = \Gamma x_t^\dagger. \tag{27} \]

The factors in this vector are all five principal components of the yields on bonds with maturities of three months, one, two, four, and five years. Their unconditional covariance matrix is the diagonal matrix of eigenvalues

\[ \text{Var}(x_t) = \Omega. \tag{28} \]
The dynamics of the rotated state vector are

\[ x_{t+1} = K x_t + \Sigma \epsilon_{t+1}, \tag{29} \]

where the parameters are defined by

\[ K = \Gamma D^\dagger \Gamma^{-1}, \quad \Sigma = \text{chol} \left( \Gamma \Sigma^\dagger \Sigma^\dagger' \Gamma' \right). \tag{30} \]

The relation between bond yields and the rotated factors is

\[ y_t = A + B x_t + \eta_t, \tag{31} \]

where the new factor loadings are

\[ B = B^\dagger \Gamma^{-1}. \tag{32} \]

These factor loadings (for all but the three-year bond yield) are the eigenvectors of the diagonalization (25).

### 3.4 Estimates of the factors’ role in the cross section

I summarize the important properties of the model rather than report parameter point estimates. This subsection focuses on the cross-sectional properties.

Table 2 describes the cross-sectional relation between the factors and bond yields. Since the factors are, by construction, principal components of yields, it is not surprising that the first few factors explain almost all of the variation in yields. We see in the first column that population standard deviations of the factors range from 5.90 for the first factor to 0.04 for the fifth. These standard errors can be thought of as the contribution of the factor to the population standard deviation of the yield curve.

The precise mapping from factors to yields is displayed in Figure 1, which plots the matrix of estimated factor loadings \( B \) scaled by the factor standard deviations. The first panel plots loadings on the first three factors. They are the usual level, slope, and curvature factors. For example, a one standard deviation increase in the first factor raises all yields by about 2.5 percentage points. The second panel plots loadings on the fourth and fifth factors. There is no obvious cross-sectional interpretation for these two factors, which appear to be economically tiny. Note the difference in scale between the two panels. A one standard deviation increase in the fifth factor does not change any yield by more than 3.5 basis points.

Because of measurement error, it is difficult to extract the final two factors from the cross-section of the term structure, even if we know the model’s parameters. Table 2 reports
the estimated standard deviation of measurement error is between five and six basis points (annualized yields). Although economically small, it is enough to obscure the effects of these factors on yields. One way to see this is to imagine a regression, using an infinite time series, of a factor on contemporaneous yields. (An econometrician cannot estimate this regression because she does not directly observe the factors.) The point estimates of the model allow analytic calculation of the $R^2$ for the regression.

The second column of statistics in Table 2 reports the $R^2$s for each factor regressed on bond yields with maturities of six months and one through five years. The effects of the first three factors on yields are sufficiently large to dominate measurement error. The $R^2$s for these factors range from 1.0 to 0.95. However, the $R^2$s for the fourth and fifth factors are only 0.62 and 0.43 respectively. Put differently, the correlations between the actual and fitted values of the factors are 0.79 and 0.66.

Kalman filtering produces more accurate estimates of the factors. Population properties of the Kalman filter are proxied by simulating one million months of bond yields (the maturities are six months and one through five years), where the “true” model is the model estimated with ML. The Kalman filter is then applied to these data, using the true parameters in the filter. The final column of Table 2 reports correlations between true and filtered estimates of the factors. These correlations are 0.84 and 0.80 for the fourth and fifth factors. Naturally, filtered estimates of the factors are more closely related to observed yields than are true factors (since observed yields are used in the filtering), as documented in the third column of statistics in Table 2.

Since only the first three factors make noticeable contributions to the cross-section of yields, why should we care about our ability to infer the other factors from the data? The reason is that the fifth factor plays an important role in yield dynamics.

### 3.5 Estimates of the factors’ role in yield dynamics

Consider investors’ $j$-month-ahead forecast of the yields used in estimation of the model. The vector of forecasts is (recall that investors know the true state vector)

$$
E_t(y_{t+j}) = A + BE_t(x_{t+j}) = A + BK^jx_t.
$$

The unconditional covariance matrix of these forecasts is

$$
\text{Var}(E_t(y_{t+j})) = BK^j\Omega B'(K^j)'.
$$

14
Because the unconditional covariance matrix of the factors $\Omega$ is diagonal, the variance in (34) can be unambiguously expressed as the sum of components attributable to each of the five factors.

Table 3 reports this decomposition. To illustrate the results, consider the first row. The table reports that twelve-month-ahead forecasts of the three-month annualized bill yield have a standard deviation of 2.23 percentage points. The vast majority of this variation is due to the first, “level” factor. The level factor has both high persistence and high unconditional volatility. The high persistence implies that the month $t$ expectation of the level factor in $t + 12$ is close to its period $t$ value. The high unconditional volatility implies that the period-$t$ level is highly variable. In combination, these effects result in substantial variation in twelve-month-ahead forecasts. More than 95 percent of the total variance of the forecast is attributable to the level factor. This pattern holds for all maturities included in the table.

The surprising result in the table is that most of the remaining variance in twelve-month-ahead forecasts is captured by the fifth factor. For each maturity, it accounts for a little more than 2.5 percent of the variance in the forecast. Although these magnitudes are small relative to the explanatory power of the level factor, they are much larger than the contribution of any other factor.

Visual evidence of the contributions of the factors to yield forecasts is in Figure 2. It displays impulse responses of the three-month bill yield to one standard deviation changes in each factor. For example, in the first panel the month-zero yield is 2.73 percentage points above its mean. Two years later, the yield remains 1.72 percentage points above its mean. The second (slope) factor corresponds to an immediate drop in the short rate of about 60 basis points, half of which has disappeared after a year. The third and fourth factors contribute little to current or future short rates. The effect of the fifth factor is qualitatively different from all of the other factors. It has no effect on the short rate at month zero. One year later, the short rate has dropped 35 basis points, where it remains for the next year. Accordingly, I label this fifth factor the “expectation” factor.

Because the expectation factor plays the central role in the remainder of the paper, it is useful to take a quick look at its time-series behavior. Figure 3 plots filtered estimates of this factor over the sample period 1964 through 2006. The factor is normalized by its model-implied population standard deviation. Its persistence is fairly low. The model’s parameter estimates imply that a shock to the factor (holding all other factors constant) has a half life of five months. Any relation between the factor and economic fluctuations is not obvious from this figure, which also displays NBER turning points. Section 4.3 uncovers a relatively high-frequency relation between the factor and economic activity.
3.6 Estimates of the factors’ role in excess return dynamics

Although the level factor is the dominant driver of yields, it plays a much less important role in expected excess returns. In this section I focus on the behavior of the log return from \( t \) to \( t + j \) on a bond with period-\( t \) maturity \( m \), in excess of the log return on a \( j \)-period bond.

The observed excess return, expressed in terms of factors and measurement error, is

\[
x_r^{(m)}_{t,t+j} = mA_m - (m-j)A_{m-j} - jA_j
+ (mB'_m - (m-j)B'_{m-j}K^j - jB'_j) x_t
- (m-j)B'_{m-j} \left( \sum_{i=1}^{j} K^{j-i} \epsilon_{t+i} \right)
+ m\eta_t^{(m)} - j\eta_t^{(m)} - (m-j)\eta_{t+j}^{(m-j)}.
\]

The four lines on the right side of (35) are, respectively, the unconditional mean, the variation in the conditional mean owing to the period-\( t \) state vector, the return innovation owing to shocks to the state vector, and the measurement error component.

The estimates of \( A \) and \( B \) allow to study directly the population properties of this excess return for a one-year horizon (\( j = 12 \)) and for bonds with maturities of two, three, four, and five years. Panel A of Table 4 reports unconditional means and standard deviations of these returns. Standard deviations are calculated for both true returns (i.e., excluding measurement error) and observed returns. The panel also reports the fraction of the total variance attributable to factor-driven variations in the conditional mean.

Unconditional mean excess annual returns are less than one percent for all of these bonds. Standard deviations of the returns range from 1.7 percent for the two-year bond to 5.5 percent for the five-year bond. We see in the panel that measurement error contributes very little to the volatility of observed returns; differences in standard deviations between true and observed returns are at most a basis point.

Panel A also reports that predictable variations in returns account for 20 to 25 percent of total return variance. Panel B decomposes this predictable variance into components attributable to each factor. Given the well-known relation between the slope of the term structure and expected excess bond returns, it is not surprising that for each bond, the “slope” factor accounts for over half of the predictable variance. A glance at Figures 1 and 2 explains why. The slope factor simultaneously raises long-term bond yields and lowers expected future short rates. The more interesting result in Panel B is that the expectation factor explains up to 30 percent of the predictable variance. Again, a glance at the two figures explains why. The expectation factor lowers expected future short rates while leaving
long-term yields unchanged.

Figure 4 displays the sensitivity of expected excess annual log returns to the level, slope, and expectation factors. For example, a one-standard-deviation increase in the slope factor raises the expected excess return to a two-year bond by about 60 basis points. The corresponding change in expected return to a five-year bond is about 200 basis points. These values are plotted with the dashed line in the figure. The solid line plots changes in expected excess returns for a one-standard deviation increase in the level factor and the dashed line plots changes for the expectation factor.

The main conclusion supported by the estimated model is that the factors that are most important for determining the shape of the term structure are not the most important in determining expected excess bond returns. Although this conclusion is consistent with the theory of Section 2.4, there are two obvious questions raised by the results. First, how can we be sure that the month-\(t\) expectation factor is truly known by investors at \(t\), instead of fit spuriously to the data by the maximum likelihood procedure? Second, is there a link between this factor and the real economy?

4 The expectations factor: additional evidence

Is the expectations factor simply a consequence of overfitting? A natural way to answer this question is to compare the factor to independent observations investors’ forecasts. At the end of the first month of every quarter since 1981Q3, participants in the Survey of Professional Forecasters are asked for their forecasts of the average level of the three-month Treasury bill during each of the next four quarters. This section examines the relation between mean forecasts (where the mean is taken across the participants) and contemporaneous values of the expectation term structure factor. Here, “contemporaneous” means the filtered estimate for the end of the first month in the quarter.

If the expectation factor is spurious, forecasters’ contemporaneous expectations should be unrelated to it. For example, assume the quarter-\(t\) level of the filtered expectation factor predicts that the short rate will decline over the next few quarters. If this prediction is simply an ex-post interpretation of the data by the maximum likelihood estimation, then the survey responses in quarter \(t\) will not anticipate a decline in rates. Thus we can test the null hypothesis that the expectation factor is entirely spurious by examining its covariation with survey forecasts of changes in rates.

Before presenting the regression results, it is instructive to study in detail two particular observations.
4.1 A tale of two Octobers

Panel A of Figure 5 displays term structures for the month-ends of October 2001 and October 2004. (The plotted points are yields for maturities of three months and one through five years.) The shapes of the term structures are similar. The three-month bill yields are both around two percent. The largest difference between the term structures is at the long end, where the October 2001 observation is 37 basis points above the October 2004 observation. The dates were chosen both because the term structures are similar and the filtered estimates of the fifth factor are not. The October 2001 estimate of this factor is about 0.67 standard deviations, while the October 2004 estimate is about −1.1 standard deviations.

This large difference in estimates of the fifth factor corresponds to a large difference in expected excess bond returns. Panel B of the figure displays model-implied expectations, as of October 2001 and October 2004, of one-year log returns to bonds in excess of the yield on a one-year bond. In 2001, the expectations are positive for all of the plotted maturities (two through five years), from 0.3 percent for the two-year bond to 1.3 percent for the five-year bond. In 2004, the expectations are negative, ranging from −0.5 percent to −1.5 percent. The difference in expected excess returns is largely accounted for by the difference in the expected time path of the three-month bill rate. Panel C reports that for 2001, the bill rate is expected to stay flat for a few months, then rise modestly. By contrast, in 2004 the bill rate is expected to rise substantially over the next year. The average difference between the two sets of forecasts over the upcoming year (November through December of the next year) is about 60 basis points.

Are these model-implied expectations reasonably consistent with investors’ expectations at the time? According to the Survey of Professional Forecasters, they are. For the surveys returned in early November 2001, the mean forecasts of the three-month bill rate for the next four quarters (2002Q1 through 2002Q4) are 1.9, 2.0, 2.4 percent, and 2.8 percent respectively. Three years later, mean forecasts are about 50 basis points higher. The forecasts for 2005Q1 through 2005Q4 are 2.3, 2.6, 2.9, and 3.2 percent. Investors (or at least those investors with beliefs similar to those embodied by the mean forecasts of the survey participants) anticipated lower expected excess returns in October 2004 than in October 2001.

Differences in expected excess returns across these two months may be related to anticipated macroeconomic activity. Forecasters responding to the 2001Q4 survey were much more pessimistic about near-term economic growth than were those responding to the 2004Q4 survey. The 2001Q1 mean forecast of real GDP growth in 2002 was 0.8 percent. By contrast, the 2004Q4 forecast of real GDP growth in 2005 was 3.5 percent. The link between the

\[1\] In particular, the months were not chosen based on the contemporaneous survey forecasts.
expectations factor and expected future economic growth is pursued in Section 4.3.

A single comparison of two months is illuminating, but not statistically compelling. The next subsection contains some regression evidence.

4.2 Regression results

Denote the quarter-\(t\) mean survey forecast of the three-month bill in quarter \(t + j\) less the quarter-\(t\) bill yield as SPF\_EXPECT\(_{t, j}\). To align the bill yield with the survey timing, the quarter-\(t\) yield is defined as the three-month yield as of the end of the first month in the quarter. The continuously compounded yield from CRSP is converted to a discount basis to match the survey. Denote quarter-\(t\) filtered estimates of the expectation factor as MODEL\_EXPECT\(_t\). These are estimates for the end of the first month in the quarter. To simplify interpretation of the estimated regression coefficients, this factor is normalized by its population standard deviation. The sample period is 1981Q3 through 2006Q4.

I first estimate the regression

\[
\text{MODEL\_EXPECT}_t = b_0 + b_1 \text{SPF\_EXPECT}(t, j) + e_{j,t}
\]  

(36)

for forecast horizons of one through four quarters \((j = 1, \ldots)\). Under the null hypothesis that the filtered estimate of the expectation factor is spurious, the coefficient \(b_1\) should be zero. Because quarterly survey forecasts are serially correlated, standard errors use the Newey-West adjustment for four lags of moving average residuals. Although the regression is probably more intuitive if the regressor and regressand are switched, there is a generated regressor problem when using the filtered estimate of the expectation factor as the explanatory variable.

The coefficient should be negative if the model’s factor is not spurious. As shown in Figure 2, the model implies that a one standard deviation increase in the expectation factor corresponds to an expected drop in the three-month bill rate of 35 basis points over the subsequent year. Reversing the order of this comparison for the purposes of (36), an expected increase in the bill rate of one percentage point corresponds to \(-2.9\) standard deviations of the factor.

Coefficient estimates for each forecast horizon are displayed in Panel A of Table 5. The null hypothesis is overwhelmingly rejected. The point estimates are reliably negative, with asymptotic \(t\) statistics ranging from \(-2.96\) to \(-5.53\). The point estimates are less than the model predicts, ranging from \(-0.5\) to \(-1.2\). In other words, the estimated factors respond less to true variations in expected changes in short rates than the model implies.

These regressions are estimated over the entire sample for which forecasts are available.
from the Survey of Professional Forecasters. From a statistical perspective, one unfortunate feature of this sample is that the estimated term structure factors are not uncorrelated. Over the entire 1964 through 2006 sample, the sample correlation between filtered values of the level and expectation factors is only 0.02. But from 1981Q3 through 2006Q4, the sample correlation is about 0.24. As Figure 2 shows, both the level and expectation factors have the same qualitative effect on expected future short rates. When the factors are high, short rates are expected to decline. Hence it is possible that the negative point estimates for (36) are proxying for the relation between the level of rates and expected future changes in rates. (Note, though, that this proxy story does not explain the tale of two Octobers.)

To control for the level of the term structure, I reverse (36) and add the estimated level factor as an additional explanatory variable. The regression is

$$\text{SPF\_EXPECT}(t, j) = b_0 + b_1 \text{MODEL\_LEVEL}_t + b_2 \text{MODEL\_EXPECT}_t + e_{j,t}. \quad (37)$$

Both explanatory variables are generated regressors. Because the expectation factor is harder to extract from the yield curve than is the level factor, there is likely to be more noise in the model’s estimate of the former factor than the latter.

Coefficient estimates for each forecast horizon are displayed in Panel B of Table 5. Both factors are negatively associated with survey expectations of future changes in the bill yield. More importantly, the statistical significance of the relation between the expectation factor and survey expectations does not disappear when the level factor is included. The asymptotic $t$ statistics for the coefficients on the expectation factor are about $-3.1$ for one-quarter-ahead and two-quarter-ahead forecasts, $-2.3$ for three-quarter-ahead forecasts, and $-1.9$ for four-quarter-ahead forecasts.

This evidence supports the model’s conclusion that the expectations factor is known by investors. In order for this factor to not affect the term structure, its predictive power for future short rates must be offset by variations in risk premia. Such a story is more plausible if the expectations factor can be linked to the business cycle.

### 4.3 The expectations factor and economic activity

I examine the lead/lag relation between filtered estimates of the expectation factor and monthly changes in log industrial production. The estimated regression is

$$100(\log(\text{IP}_t) - \log(\text{IP}_{t-1})) = b_{0,i} + b_{1,i} \text{MODEL\_EXPECT}_{t-i} + e_{t,i}, \quad i = -6, \ldots, 6. \quad (38)$$
The change in IP lags the expectation factor for $i < 0$ and leads it for $i > 0$. Log changes in IP are serially correlated. A typical serial correlation of fitted residuals for (38) is about 0.3. I therefore report Newey-West standard errors adjusted for two lags of moving average residuals. As in Section 4.2, the expectations factor is normalized by its population standard deviation.

Estimation results are in Table 6. There is a strong, statistically significant inverse relation between industrial production and the expectations factor. In other words, low growth in industrial production corresponds to high risk premia accompanied by expected future declines in short-term rates. Growth in industrial production begins to drop a few months prior to the increase in the expectations factor, continuing for a couple of months after the increase in the expectations factor. If the expectations factor is a standard deviation above its mean in month $t$, monthly growth in industrial production in months $t - 5$ through $t + 2$ averages about 11 basis points per month below average. (To put the 11 basis points in perspective, the standard deviation of monthly IP growth is about 70 basis points.)

These results are comforting because they are qualitatively consistent with a simple story. Investors believe that the Fed will attempt to offset some types of short-lived macroeconomic shocks with monetary policy actions. The Fed action is not anticipated to be immediate; short rates may not change for a number of months. The same macroeconomic shocks change investors’ willingness to bear risk. Thus the net effect of the macro shocks on current yields is muted because the expected change in short rates and the change in risk premia work in opposite directions.

5 Forecasting with the cross section

The empirical evidence in the previous section tells us there is a term structure factor that can predict future yields and excess bond returns, yet is difficult to detect in the cross-section. This conclusion leads to a natural question. If we follow standard practice by forecasting excess returns using only information in the cross-section of yields, how accurate will we be?

According to the estimated model (and as summarized in Table 4), between 22 and 25 percent of the total variance in annual excess log returns to bonds corresponds to variation in conditional mean returns. An econometrician cannot reproduce a time series of conditional expected returns, even given an infinite time series, because she must filter the factors instead of observe them. This section considers how accurately she estimates expected excess using standard predictive regressions instead of filtering. The forecasting variables are taken from cross-section of the term structure.

Following Cochrane and Piazzesi (2005) (hereafter CP), the regressions predict, as of
month $t$, the excess log return to a bond from month $t$ to month $t + 12$. One regression is inspired by standard three-factor term structure models. It uses month-$t$ values of the level, slope, and curvature of the term structure. They are respectively defined as the five-year yield, five-year yield less three-month yield, and two-year yield less the average of the three-month and five-year yields. The other regression uses the five forward rates that CP found to contain substantial information about future excess returns. They are the month-$t$ forward rates from year $m$ to year $m + 1$ for $m = 0, \ldots, 4$.

Population and finite-sample properties of the regressions are calculated by assuming that the term structure model estimated in Section 3 is correct. Population properties are determined analytically, while finite-sample properties are produced with Monte Carlo simulations. The length of each simulated sample is 516 months (43 years), which is the length of the sample used to estimate the model in Section 3. The results here are based on 10,000 simulations.

We first take a close look at predictions of annual excess returns to the five-year bond. The relevant information is in Table 7. Panel A contains results for the level, slope, and curvature regression, while Panel B contains results for the forward-rate regression. The population $R^2$s of the two regressions are similar; 16 percent using level, slope, and curvature, and 17 percent using five forward rates. For comparison, Table 4 reports that 23 percent of the excess return variance is truly predictable. Because factors cannot be precisely inferred from the term structure, the regressions do not capture everything that investors know, even with an infinite time series. The population coefficients on the five forward rates display the tent shape uncovered by CP. However, these coefficients are closer to zero than are those estimated by CP. Similarly, the population $R^2$ here is less than the $R^2$ of 36 percent estimated by CP.

These differences in magnitudes are probably a consequence of model misspecification. The sample period used to estimate the model in Section 3 is almost identical to that used by CP. Thus one way to interpret these differences is that the model does not reproduce the strong annual-horizon predictability that is a feature of the data used to estimate the model. However, there is relatively little information in this estimated predictability. The finite-sample evidence in Table 7 shows that both parameter estimates and $R^2$s have very large sampling uncertainty. For example, the 95 percent confidence range for the forward-rate regression’s $R^2$ is from 9 percent to 40 percent. The range for the other regression’s $R^2$ is similar. In a given sample, the two regressions produce substantially different $R^2$s. The evidence is in Panel C of Table 7. The mean and median difference is only one percentage point, but a 95 percent confidence interval for the difference ranges from $-0.043$ (i.e., the three-factor $R^2$ is 4.3 percentage points greater than the forward-rate $R^2$) to 8.2%. 

22
Moreover, although the tent shape is a population property of the model, it is not a result that leaps out of a random sample of data generated by the model. In 40 percent of these finite samples, the forward-rate regression parameter estimates do not satisfy the tent shape.² By contrast, the coefficients on level and slope for the other regression are almost always positive, as indicated by their 95 percent confidence bounds in the table.

In population, the forward-rate regression captures slightly more of the true predictability of excess returns than does the three-factor regression. Is this also true for the sample size studied here? More precisely, is the average difference between the fitted prediction and the true expectation smaller for the forward-rate regression than for the three-factor regression? To answer this question, I use a root mean squared error metric with the following notation. Individual simulations are indexed by \(i = 1, \ldots, 10,000\). Annual log excess returns are predicted with OLS regressions for maturities of one through five years. Maturities are measured in months by \(m\). For simulation \(i\), define the root mean squared error for an \(m\)-maturity bond as

\[
RMSE(i, m, p) = \sqrt{\frac{1}{516 - 12} \sum_{t=1}^{516-12} \left( E_t(x_{m,t+t+12}) - E_t^{[p]}(x_{m,t+t+12}) \right)^2}.
\] (39)

The first term on the right of (39) is the true conditional expected excess return to an \(m\)-month bond from \(t\) to \(t + 12\). The second term is the fitted in-sample expectation from a predictive regression. In-sample forecasts for the three-factor regression are indicated with \(p = 1\) and the corresponding forecasts for the forward-rate regression are indicated with \(p = 2\). Note that this forecast error measures the distance between two forecasts, not the difference between forecasts and realizations.

Table 8 contains the relevant results. The first column of statistics contains the population standard deviations of conditional expected annual returns. These standard deviations are useful benchmarks for the reported RMSEs because they can also be interpreted as RMSEs. Consider constant forecasts of annual excess returns equal to the unconditional mean excess returns. The RMSE of this forecast (again, defining RMSE relative to the true conditional forecast, not relative to realized returns) equals the standard deviation of the conditional expected excess return.

According to the RMSE metric, both regressions outperform the constant-expectation benchmark, and the forward-rate regression outperforms the three-factor regression. Consider, for example, forecasts of annual excess returns to a five-year bond. The population

²This result is not found in any table. The estimates are defined to satisfy the tent shape if the five estimated coefficients satisfy \(1^{st} < 2^{nd} < 3^{rd} > 4^{th} > 5^{th}\).
standard deviation of expected excess returns is 2.62 percent. The mean RMSEs for the three-factor and forward-rate regressions are 1.95 percent and 1.84 percent respectively. The three-factor RMSE exceeds the forward-rate RMSE in 84 percent of the individual simulations. The straightforward conclusion to draw from these results is that the additional information in two additional points on the term structure outweighs the problem of overfitting, at least for samples of the size studied here.

6 Conclusion

This paper shows that an econometrician cannot extract from bond yields all information investors have about expected future yields. An “expectation” factor contains information about expected future yields but is hidden, in the sense that it has a negligible effect on the term structure. Estimation procedures that explicitly look for a hidden factor, such as filtering, are helpful, but are no substitute for direct observation. One lesson to draw from these results is that information from sources other than bond yields can be valuable in uncovering hidden factors. The evidence here shows that the expectation factor is related to both real activity and investors’ responses to surveys.

Moreover, nothing in the theory implies that hidden factor for one type of asset must also be hidden from the perspective of other types of assets. An important question—but one that is outside the scope of this paper—is whether information from stock and foreign exchange markets can be used to build more accurate term structure models.
References


Table 1. Summary statistics for Treasury yields, 1964 through 2006.

Month-end yields on zero-coupon Treasury bonds are from CRSP. The date range is January 1964 through December 2006. Yields are continuously compounded and expressed in percent per year.

Panel A. Univariate statistics

<table>
<thead>
<tr>
<th>Maturity</th>
<th>3 mon</th>
<th>1 yr</th>
<th>2 yr</th>
<th>3 yr</th>
<th>4 yr</th>
<th>5 yr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Std dev</td>
<td>2.785</td>
<td>2.751</td>
<td>2.668</td>
<td>2.859</td>
<td>2.523</td>
<td>2.481</td>
</tr>
</tbody>
</table>

Panel B. First five principal components of the unconditional covariance matrix

<table>
<thead>
<tr>
<th>Index of component</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eigenvalue</td>
<td>40.6264</td>
<td>0.9195</td>
<td>0.0689</td>
<td>0.0083</td>
<td>0.0051</td>
</tr>
<tr>
<td>Eigenvectors</td>
<td>3 mon</td>
<td>0.424</td>
<td>-0.673</td>
<td>0.593</td>
<td>-0.123</td>
</tr>
<tr>
<td></td>
<td>1 yr</td>
<td>0.428</td>
<td>-0.347</td>
<td>-0.578</td>
<td>0.590</td>
</tr>
<tr>
<td></td>
<td>2 yr</td>
<td>0.418</td>
<td>0.019</td>
<td>-0.374</td>
<td>-0.560</td>
</tr>
<tr>
<td></td>
<td>3 yr</td>
<td>0.403</td>
<td>0.235</td>
<td>-0.129</td>
<td>-0.336</td>
</tr>
<tr>
<td></td>
<td>4 yr</td>
<td>0.392</td>
<td>0.382</td>
<td>0.172</td>
<td>-0.015</td>
</tr>
<tr>
<td></td>
<td>5 yr</td>
<td>0.382</td>
<td>0.475</td>
<td>0.358</td>
<td>0.458</td>
</tr>
</tbody>
</table>
Table 2. Model-implied population properties of term structure factors

A five-factor Gaussian term structure model is estimated with the Kalman filter. True yields are affine functions of the unobserved factors. Observed yields are contaminated with iid measurement error. The data are month-end yields, from January 1964 through December 2006, on zero-coupon bonds with maturities of three months and one through five years. The factors are rotated to represent, in order, the first five principal components of the bond yields (expressed in percent per year). The first column of the table reports the population standard deviations of the factors and the measurement error. The second column reports the population $R^2$ of a theoretical contemporaneous regression of the true, unobserved factors on all six observed bond yields. The third column reports the population $R^2$ of similar regressions using filtered estimates of the factors in place of the true factors. The fourth column reports correlations between true factors and filtered estimates of the factors.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Std dev</th>
<th>$R^2$'s of factors on yields</th>
<th>Correl of true, filtered factor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>True factors</td>
<td>Filtered factors</td>
</tr>
<tr>
<td>1</td>
<td>5.901</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>2</td>
<td>0.929</td>
<td>0.997</td>
<td>1.000</td>
</tr>
<tr>
<td>3</td>
<td>0.254</td>
<td>0.945</td>
<td>0.993</td>
</tr>
<tr>
<td>4</td>
<td>0.067</td>
<td>0.619</td>
<td>0.876</td>
</tr>
<tr>
<td>5</td>
<td>0.043</td>
<td>0.434</td>
<td>0.684</td>
</tr>
<tr>
<td>Measurement error</td>
<td>0.0558</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3. Decomposition of variances of 12-month-ahead yield forecasts

A five-factor Gaussian term structure model is estimated with the Kalman filter. The factors represent, in order, the first five principal components of the bond yields and are unconditionally uncorrelated. Parameter estimates are used to construct unconditional variances of 12-month-ahead expectations of bond yields. The table reports decompositions of these variances into components attributable to each of the five factors.

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Std dev of forecast (%/year)</th>
<th>Fract of var attributable to factor 1</th>
<th>Fract of var attributable to factor 2</th>
<th>Fract of var attributable to factor 3</th>
<th>Fract of var attributable to factor 4</th>
<th>Fract of var attributable to factor 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 mon</td>
<td>2.23</td>
<td>0.955</td>
<td>0.018</td>
<td>0.001</td>
<td>0.001</td>
<td>0.026</td>
</tr>
<tr>
<td>1 yr</td>
<td>2.25</td>
<td>0.966</td>
<td>0.006</td>
<td>0.000</td>
<td>0.001</td>
<td>0.027</td>
</tr>
<tr>
<td>2 yr</td>
<td>2.28</td>
<td>0.971</td>
<td>0.000</td>
<td>0.001</td>
<td>0.000</td>
<td>0.028</td>
</tr>
<tr>
<td>3 yr</td>
<td>2.25</td>
<td>0.967</td>
<td>0.002</td>
<td>0.004</td>
<td>0.000</td>
<td>0.027</td>
</tr>
<tr>
<td>4 yr</td>
<td>2.25</td>
<td>0.959</td>
<td>0.006</td>
<td>0.008</td>
<td>0.000</td>
<td>0.027</td>
</tr>
<tr>
<td>5 yr</td>
<td>2.23</td>
<td>0.949</td>
<td>0.011</td>
<td>0.011</td>
<td>0.000</td>
<td>0.028</td>
</tr>
</tbody>
</table>
Table 4. Model-implied properties of annual excess bond returns

A five-factor Gaussian term structure model is estimated with the Kalman filter. The factors represent, in order, the first five principal components of the bond yields and are unconditionally uncorrelated. Parameter estimates are used to calculate population properties of annual log returns to bonds in excess of the log return to a one-year bond. In Panel A, return variances are calculated for both true excess returns and observed excess returns. The latter are contaminated by measurement error. The columns labeled “Predictable frac of var” report the fraction of the variance attributable to time-variation in conditional means of true returns. Panel B reports variance decompositions of true conditional expected excess returns.

Panel A. Univariate statistics

<table>
<thead>
<tr>
<th>Maturity</th>
<th>True returns</th>
<th>Observed returns</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std dev</td>
</tr>
<tr>
<td>2 yr</td>
<td>0.37</td>
<td>1.69</td>
</tr>
<tr>
<td>3 yr</td>
<td>0.68</td>
<td>3.13</td>
</tr>
<tr>
<td>4 yr</td>
<td>0.86</td>
<td>4.40</td>
</tr>
<tr>
<td>5 yr</td>
<td>0.86</td>
<td>5.51</td>
</tr>
</tbody>
</table>

Panel B. Decomposition of variances of expected excess returns

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Fract of var attributable to factor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>2 yr</td>
<td>0.277</td>
</tr>
<tr>
<td>3 yr</td>
<td>0.145</td>
</tr>
<tr>
<td>4 yr</td>
<td>0.118</td>
</tr>
<tr>
<td>5 yr</td>
<td>0.102</td>
</tr>
</tbody>
</table>
Table 5. Model-implied expectations compared to survey forecasts

Quarterly observations of expectations of future Treasury bill yields are from the Survey of Professional Forecasters. The data used are quarter-\( t \) mean survey forecasts of the three-month T-bill yield during quarters \( t + j, j = 1, \ldots 4 \). The contemporaneous three-month yield is subtracted from the forecasts to produce forecasted changes in the yield. Contemporaneous filtered estimates of the “level” and “expectation” factors are taken from a five-factor term structure model. The factors are normalized to have standard deviations of one. All regressions are estimated from 1981Q3 through 2006Q4 (102 quarterly observations). Newey-West standard errors are in parentheses, adjusted for four lags of moving average residuals.

Panel A. Regressions of the expectation factor on the survey-based expected change

<table>
<thead>
<tr>
<th>Quarters ahead ((j))</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coef</td>
<td>-1.162</td>
<td>-0.858</td>
<td>-0.627</td>
<td>-0.483</td>
</tr>
<tr>
<td></td>
<td>(0.210)</td>
<td>(0.211)</td>
<td>(0.193)</td>
<td>(0.164)</td>
</tr>
<tr>
<td>AR(1) of residual</td>
<td>0.54</td>
<td>0.57</td>
<td>0.57</td>
<td>0.58</td>
</tr>
</tbody>
</table>

Panel B. Regressions of the survey-based expected change on the level and expectation factors

<table>
<thead>
<tr>
<th>Quarters ahead ((j))</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coef on level</td>
<td>-0.106</td>
<td>-0.147</td>
<td>-0.207</td>
<td>-0.268</td>
</tr>
<tr>
<td></td>
<td>(0.036)</td>
<td>(0.046)</td>
<td>(0.058)</td>
<td>(0.073)</td>
</tr>
<tr>
<td>Coef on expectation</td>
<td>-0.134</td>
<td>-0.152</td>
<td>-0.145</td>
<td>-0.153</td>
</tr>
<tr>
<td></td>
<td>(0.044)</td>
<td>(0.051)</td>
<td>(0.062)</td>
<td>(0.082)</td>
</tr>
<tr>
<td>AR(1) of residual</td>
<td>0.22</td>
<td>0.46</td>
<td>0.52</td>
<td>0.55</td>
</tr>
</tbody>
</table>
Table 6. The relation between industrial production and the expectations factor

The log change industrial production from month $t - 1$ to month $t$ is regressed on the month $t - i$ filtered estimate of the expectations factor, for $i = -6, \ldots, 6$. The log change is expressed in percent and the factor is normalized to have a standard deviation of one. Newey-West standard errors are calculated using two lags of moving average residuals. The sample period is 1964 through 2006.

<table>
<thead>
<tr>
<th>Lead of $\Delta \log(IP)$</th>
<th>Coef</th>
<th>Std error</th>
<th>$t$-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>−6</td>
<td>−0.027</td>
<td>0.048</td>
<td>−0.55</td>
</tr>
<tr>
<td>−5</td>
<td>−0.098</td>
<td>0.050</td>
<td>−1.96</td>
</tr>
<tr>
<td>−4</td>
<td>−0.101</td>
<td>0.051</td>
<td>−1.98</td>
</tr>
<tr>
<td>−3</td>
<td>−0.132</td>
<td>0.050</td>
<td>−2.63</td>
</tr>
<tr>
<td>−2</td>
<td>−0.145</td>
<td>0.049</td>
<td>−2.95</td>
</tr>
<tr>
<td>−1</td>
<td>−0.120</td>
<td>0.050</td>
<td>−2.40</td>
</tr>
<tr>
<td>0</td>
<td>−0.137</td>
<td>0.054</td>
<td>−2.51</td>
</tr>
<tr>
<td>1</td>
<td>−0.113</td>
<td>0.057</td>
<td>−1.97</td>
</tr>
<tr>
<td>2</td>
<td>−0.089</td>
<td>0.054</td>
<td>−1.65</td>
</tr>
<tr>
<td>3</td>
<td>−0.060</td>
<td>0.047</td>
<td>−1.26</td>
</tr>
<tr>
<td>4</td>
<td>−0.071</td>
<td>0.049</td>
<td>−1.44</td>
</tr>
<tr>
<td>5</td>
<td>−0.038</td>
<td>0.052</td>
<td>−0.73</td>
</tr>
<tr>
<td>6</td>
<td>−0.078</td>
<td>0.051</td>
<td>−1.52</td>
</tr>
</tbody>
</table>
Table 7. Population and finite-sample properties of predictive regressions

Excess log returns to a five-year bond from month $t$ to month $t + 12$ are predicted with the month-$t$ shape of the term structure using two OLS regressions. The first regression uses the level, slope, and curvature of the term structure, as defined in the text. The second regression uses five forward rates. The true data-generating process is this paper’s estimated term structure model. Population values of the coefficients and $R^2$'s are calculated analytically. Finite-sample properties use simulations of 516 months of bond yields. The table summarizes results from 10,000 simulations. The notation $F(m, n)$ denotes the forward rate from year $m$ to year $n$.

Panel A. Predicting excess returns with level, slope, and curvature

<table>
<thead>
<tr>
<th>Coefficient on:</th>
<th>Level</th>
<th>Slope</th>
<th>Curvature</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population value</td>
<td>0.471</td>
<td>1.575</td>
<td>0.634</td>
<td>0.16</td>
</tr>
<tr>
<td>Mean across sims</td>
<td>0.894</td>
<td>1.684</td>
<td>0.198</td>
<td>0.22</td>
</tr>
<tr>
<td>Std dev across sims</td>
<td>0.487</td>
<td>0.780</td>
<td>2.812</td>
<td>0.08</td>
</tr>
<tr>
<td>95 percent interval</td>
<td>[0.15 2.04]</td>
<td>[0.07 3.15]</td>
<td>[−5.36 5.62]</td>
<td>[0.07 0.39]</td>
</tr>
</tbody>
</table>

Panel B. Predicting excess returns with five forward rates

<table>
<thead>
<tr>
<th>Coefficient on:</th>
<th>$F(0, 1)$</th>
<th>$F(1, 2)$</th>
<th>$F(2, 3)$</th>
<th>$F(3, 4)$</th>
<th>$F(4, 5)$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population value</td>
<td>−1.982</td>
<td>0.631</td>
<td>2.687</td>
<td>0.272</td>
<td>−1.108</td>
<td>0.17</td>
</tr>
<tr>
<td>Mean across sims</td>
<td>−1.885</td>
<td>0.588</td>
<td>2.644</td>
<td>0.441</td>
<td>−0.868</td>
<td>0.23</td>
</tr>
<tr>
<td>Std dev across sims</td>
<td>0.962</td>
<td>1.661</td>
<td>1.053</td>
<td>0.853</td>
<td>0.872</td>
<td>0.08</td>
</tr>
<tr>
<td>95 percent interval</td>
<td>[−3.76 0.03]</td>
<td>[−2.70 3.77]</td>
<td>[0.62 4.77]</td>
<td>[−1.23 2.11]</td>
<td>[−2.60 0.86]</td>
<td>[0.09 0.40]</td>
</tr>
</tbody>
</table>

Panel C. Probability distribution of the difference in finite-sample $R^2$'s

<table>
<thead>
<tr>
<th>Percentile</th>
<th>2.5%</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>97.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forward-rate $R^2$ less 3-factor $R^2$</td>
<td>−0.043</td>
<td>−0.002</td>
<td>0.014</td>
<td>0.032</td>
<td>0.082</td>
</tr>
</tbody>
</table>
Table 8. Finite-sample forecast accuracy of predictive regressions

Excess log returns to $m$-year bonds from month $t$ to month $t+12$, $m = 1, \ldots, 5$, are predicted with the month-$t$ shape of the term structure using two OLS regressions. Regression [1] uses level, slope, and curvature. Regression [2] uses five forward rates. The true data-generating process is this paper’s estimated term structure model. Finite-sample properties use simulations of 516 months of bond yields. The table summarizes monthly differences between fitted month-$t$ forecasts and true month-$t$ expectations of expected excess returns. For each simulation, the square root of the mean squared difference, denoted RMSE, is calculated for each regression. The table summarizes results from 10,000 simulations. All values are in percent per year.

<table>
<thead>
<tr>
<th>Bond maturity</th>
<th>Std dev of true expectation</th>
<th>Mean RMSE across simulations</th>
<th>Fraction of sims with RMSE[1] &gt; RMSE[2]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Std dev</td>
<td>Mean RMSE</td>
<td></td>
</tr>
<tr>
<td>2 years</td>
<td>0.803</td>
<td>0.576</td>
<td>0.541</td>
</tr>
<tr>
<td>3 years</td>
<td>1.497</td>
<td>1.173</td>
<td>1.078</td>
</tr>
<tr>
<td>4 years</td>
<td>2.177</td>
<td>1.675</td>
<td>1.522</td>
</tr>
<tr>
<td>5 years</td>
<td>2.616</td>
<td>1.948</td>
<td>1.839</td>
</tr>
</tbody>
</table>
Fig. 1. Estimated loadings of yields on the five factors of a term structure model. Each line represents the response of the term structure to a one standard deviation variation in the given factor.
Fig. 2. Responses of the three-month bill rate to term structure factors. Each panel plots the expected time path of the three-month bill yield, assuming that at month zero the specified factor is one standard deviation above its mean. All other factors are set to their unconditional means.
Fig. 3. Filtered estimates of the “expectation” factor. The vertical lines are NBER business cycle break points.
Fig. 4. Sensitivity of expected excess bond returns to term structure factors. The month-{:t} expected annual log return to a m-year bond less the log return to a one-year bond depends on the month-{:t} values of the term structure factors. The figure plots, for m = 2 through m = 5, the sensitivity of the expected excess return to one-standard-deviation changes in the “level” factor (solid line), the “slope” factor (dashed line), and “expectation” factor (dotted line).
Fig. 5. A comparison of October 2001 and October 2004. Values for the two months are plotted with '+’ and ‘o’ respectively. Panel A displays the month-end term structures. Panel B displays model-implied expected excess log returns (over the one-year yield) for bonds with maturities of two through five years. Panel C displays expected future three-month yields over the next 24 months, where month zero is October of 2001 and 2004 respectively. Panel D displays expected future five-year yields.