Sharpe ratios in term structure models

Gregory R. Duffee[*]

Johns Hopkins University

First draft November 2009

This Draft: April 7, 2010

ABSTRACT

Conditional maximum Sharpe ratios implied by fully flexible four-factor and five-factor Gaussian term structure models are astronomically high. Estimation of term structure models subject to a constraint on their Sharpe ratios uncovers properties that hold for a wide range of Sharpe ratios. These robust properties include (a) an inverse relation between a bond's maturity and its average Sharpe ratio; (b) between 15 and 20 percent of annual excess returns to bonds are predictable; and (c) variations in expected excess bond returns are driven by two factors. These factors operate at different frequencies. Nonrobust features include the mean level of the term structure. Unconstrained models imply that investors anticipated much of the decline of interest rates in the 1990s. Constrained models disagree.

[*]Voice 410-516-8828, email duffee@jhu.edu. See www.econ.jhu.edu/people/Duffee for the latest version.
Thanks to John Cochrane and seminar participants at LSE for helpful comments.

1 Introduction

How predictable are expected excess returns to Treasury bonds? Attempts to answer this question have evolved from the univariate regression approach of Fama and Bliss (1987) to high-dimensional predictive regressions and forecasts from multifactor no-arbitrage dynamic term structure models. These more flexible tools have uncovered substantial in-sample variability of conditional expected excess returns.

Flexibility and overfitting go hand-in-hand. Estimated models may be uncovering sample-specific patterns instead of features of the true data-generating process. A variety of standard econometric methods are used to analyze and adjust for overfitting, such as studying out-of-sample properties. In addition, the built-in flexibility of dynamic term structure models is often pared down by imposing plausible parametric restrictions. But the most commonly applied check on overfitting is intuition. Is the estimated predictability unbelievably large? This metric is typically applied informally; in discussions, at conferences, and in the mind of the researcher.

This paper puts some formal structure on our intuition through the use of conditional Sharpe ratios. No-arbitrage term structure models specify the dynamics of the stochastic discount factor used to price fixed-income instruments. Armed with estimates of a model's parameters, we can construct maximum conditional Sharpe ratios, as well as conditional Sharpe ratios for arbitrary bond portfolios. Large Sharpe ratios are evidence of overfitting, thus the ratios can be used as an informal specification test. Similarly, more realistic models can be estimated by imposing a constraint on allowable Sharpe ratios.

There are two clear limitations of this approach. First, plausible bounds on Sharpe ratios are unavoidably subjective. As researchers have uncovered ever more profitable investment strategies, acceptable bounds have risen. Ross (1976) uses a maximum annual Sharpe ratio of about 0.25. MacKinlay (1995) uses 0.6, while Cochrane and Saa-Requejo (2000) use 1.0. Second, existing dynamic term structure models are insufficiently flexible to capture the empirical dynamics of both conditional means and conditional covariances. Thus either the numerator or denominator of the conditional Sharpe ratio is likely misspecified.

Faced with the choice of fitting either first or second moments, I follow much of the recent literature by examining Gaussian models. This choices maximizes the flexibility of conditional means. Gaussian models also give us a reasonably clean interpretation of the model-implied conditional Sharpe ratios. Because the models rule out time-variation in conditional covariances, model-implied Sharpe ratios are best thought of as expected excess returns divided by average standard deviations rather than by conditional standard deviations.

Owing to both of these limitations, it makes little sense to argue that a particular estimated model is better than another because its maximum conditional annual Sharpe ratios are around, say, 1.0 instead of 2.0. In this paper I do not take (much of) a stand on the maximum plausible Sharpe ratio. Instead, I focus on two issues. First, I examine how the properties of estimated models vary as a constraint on maximum Sharpe ratios varies. What features of these models are robust to the level of the constraint, and which are highly sensitive? Second, I examine how a model with n factors compares to a model with $n+1$ factors, holding constant a constraint on maximum conditional Sharpe ratios. Are the models effectively similar when the constraint is imposed?

I estimate the models using artificial zero-coupon Treasury bond yields that are assumed to be observed with noise. The data are monthly from 1971 through 2008. The likelihood function of the Kalman filter is maximized with and without constraining conditional Sharpe ratios. One clear result is that Sharpe ratios implied by unconstrained, high-dimensional Gaussian models are much too high. The magnitudes involved are stunning. For example, using simple returns, the sample mean of conditional maximum Sharpe ratios produced by an estimated five-factor model is on the order of 10^{30}. Such an extreme value is not just an artifact of particular data and model used here. The no-arbitrage model embedded in the estimates of Cochrane and Piazzesi (2005) produces a corresponding value on the order of 10^{18}.

Unless they are constrained in some way, estimated models with more than three factors are inconsistent with anyone's view of sensible Sharpe ratio bounds. At least qualitatively, this conclusion is not surprising to researchers active in term structure estimation. Unconstrained high-dimensional models are used only to illustrate specific points, such as the possible role of hidden factors explored in Duffee (2008) and the construction by Cochrane and Piazzesi (2005) of a model that reproduces arbitrary VAR dynamics. Models designed to be taken more seriously are restricted in a variety of ways. For example, Duffee (2002) sets to zero many parameters that are statistically insignificant. Joslin, Priebsch, and Singleton (2009) and Cochrane and Piazzesi (2008) adopt maintained assumptions about the number of factors that are allowed to drive variations in risk premia.

The methodology here imposes no parameter-specific restrictions. Only Sharpe ratio constraints are imposed. This approach reveals some important features of the data sample that are not sensitive to the Sharpe constraint. First, unconditional Sharpe ratios for bonds are inversely related to maturity. For short maturities (say, three or six months), these ratios are in the neighborhood of the unconditional Sharpe ratio for the aggregate stock market. The two features of the data driving this inverse relation are (a) the average slope of the term structure is steeper at the short end; and (b) volatilities of yields vary little across maturities. Dynamic term structure models attribute this pattern to risk premia on "level" and "slope" risk. Investors are compensated for the risk that the term structure jumps up; all bonds face this risk. Investors are also compensated for the risk that the slope of the term structure falls. Long-maturity bonds hedge this risk, while short-maturity bonds are exposed to this risk.

Another feature that is robust to the constraint on Sharpe ratios is that more than one factor drives variations in expected excess bond returns. More precisely, we can construct a single factor that accounts for almost all the variation in expected excess monthly returns, and a single factor that accounts for almost all the variation in expected excess annual returns, but the factors are not the same. Instead, there are both high-frequency and low-frequency variations in expected excess returns. Around 15 to 20 percent of annual excess bond returns are predictable. Of this predictable variation, roughly 30 percent is orthogonal to the factor that explains more than 99 percent of the variation in monthly excess returns.

A constraint on maximum Sharpe ratios has large effects on two features of the term structure (other than on the Sharpe ratios themselves). First, tighter Sharpe ratio bounds correspond to higher unconditional mean yield curves. Unconstrained models imply that investors anticipated much of the decline in bond yields from the end of 1988 through the end of 2000; yields at year-end 1988 were substantially above their unconditional means. For reasonable Sharpe ratios, little to none of this drop was expected. Second, tighter Sharpe ratio bounds reduce the predictability of excess returns to short-maturity bonds substantially, and much more than the bounds reduce the corresponding predictability for long-maturity bonds.

Using a root mean squared error metric, in-sample accuracy varies little across models with three to five factors, regardless of the constraint imposed on Sharpe ratios. When Sharpe ratios are constrained to be equal across models with different numbers of factors, neither cross-sectional accuracy nor forecast accuracy hinges on the dimension of the model. Differences in cross-sectional root mean squared errors are at most about three basis points of annualized yields, while differences in forecast accuracy at a twelve-month horizon are around one to three basis points.

The next section briefly notes earlier uses of Sharpe ratios in model evaluation and estimation. Sharpe ratio mathematics is reviewed in Section 3 and the term structure setting is outlined in Section 4. Empirical evidence is contained in Sections 5 and 6. The final section concludes.

2 Some earlier literature on Sharpe ratios

The first use of Sharpe ratios in model evaluation is Ross (1976), ten years after Sharpe (1966) introduced the concept. He imposes a subjective maximum Sharpe ratio on asset portfolios to estimate deviations from arbitrage pricing theory. MacKinlay (1995) uses a subjective bound on maximum Sharpe ratios as a specification test of the Fama-French model. In a no-arbitrage setting, Hansen and Jagannathan (1991) develop the link between maximum Sharpe ratios and the volatility of stochastic discount factors. Hansen and Jagannathan are also the first to use properties of returns to Treasury securities to estimate the volatility of the SDF. For maturities from three to twelve months, they report unconditional Sharpe ratios in the neighborhood of one at a quarterly horizon.[1]

Cochrane and Saa-Requejo (2000) use the link between Sharpe ratios and stochastic discount factors to place "good-deal" bounds on the prices of derivative securities. They take the physical dynamics of various stochastic processes as given and use a bound on maximum conditional Sharpe ratios to place plausible bounds on the allowable equivalent-martingale dynamics. By constrast, my approach here is to find the best combination, in a likelihood sense, of physical and equivalent-martingale dynamics subject to a bound on maximum conditional Sharpe ratios.

Sangvinatsos and Wachter (2005) are probably the first to analyze conditional Sharpe ratios in dynamic term structure models. To study dynamic portfolio choice, they construct a time series of conditional maximum Sharpe ratios implied by a parameterized model and its factor realizations. They use the results to examine the variation in investment opportunities. I construct similar time series of Sharpe ratios but use them in model evaluation.

3 Standard Sharpe ratio mathematics

In a typical paper that uses model-implied Sharpe ratios, the discussion of Sharpe ratio mathematics takes one or two paragraphs. Unfortunately, I must devote many pages to the subject. The main reason for this in-depth analysis is that I calculate Sharpe ratios for returns measured over discrete horizons. Discrete horizons are necessary to line up a model's results with empirically-observed Sharpe ratios. However, they also drive a wedge between Sharpe ratios calculated using simple returns and those calculated using log returns.

The size of this wedge depends on the conditional volatility of the stochastic discount factor (SDF). For what we could call "realistic" conditional volatilities, the wedge is unim-

[1] They also study properties of shorter-maturity bills, but those results are not directly applicable to the empirical analysis here.

portant. But for the conditional volatilities that are implied by some of the models estimated here, the wedge is, to put it mildly, large. In addition, as we will see, the conditional correlation between discrete-horizon simple bond returns and the stochastic discount factor *decreases* in the conditional volatility of the SDF. Thus when the volatility of the SDF is high, conditional Sharpe ratios of bond portfolios need not be close to maximum conditional Sharpe ratios that can be attained with other fixed-income instruments.

3.1 Sharpe ratios using simple excess returns

Financial instrument i, which may be a portfolio or investment strategy, has a period-t value $P_{i,t}$ and a payoff next period that is the sum of a cash flow $D_{i,t+1}$ and an ex-dividend value $P_{i,t+1}$. If the period-t price is nonzero, define the gross simple return as

$$R_{i,t+1} \equiv \frac{P_{i,t+1} + D_{i,t+1}}{P_{i,t}}, \quad P_{i,t} \neq 0. \tag{1}$$

Assume there is a one-period riskless bond with gross return in period $t+1$ of $R_{f,t+1}$. Define the excess simple return to portfolio i as

$$R^e_{i,t+1} = \begin{cases} R_{i,t+1} - R_{f,t+1}, & P_{i,t} \neq 0; \\ D_{i,t+1} + P_{i,t+1}, & P_{i,t} = 0. \end{cases} \tag{2}$$

Define instrument i's Sharpe ratio using simple returns as

$$S^i_t \equiv \frac{E_t R^e_{i,t+1}}{\sqrt{\mathrm{Var}_t\left(R^e_{i,t+1}\right)}}. \tag{3}$$

The upper-case S denotes a Sharpe ratio using simple returns. Ratios for log returns, defined in Section 3.2, use a lower-case s.

The law of one price implies the existence of a possibly unique stochastic discount factor (SDF) $M_{t+1} = \pi_{t+1}/\pi_t$ such that

$$0 = E_t\left(R^e_{i,t+1} M_{t+1}\right). \tag{4}$$

Similarly, the gross return to the riskless bond satisfies

$$R_{f,t+1} \equiv \exp\left(r_t\right) = \frac{1}{E_t(M_{t+1})}, \tag{5}$$

where r_t is the riskless bond's continuously compounded yield. Then (4) and (5) can be used to express the Sharpe ratio of instrument i as

$$S_t^i = -R_{f,t+1}\mathrm{Cor}_t\left(R_{i,t+1}^e, M_{t+1}\right)\sqrt{\mathrm{Var}_t\left(M_{t+1}\right)}. \tag{6}$$

An upper-case Θ denotes the maximum Sharpe ratio using simple returns,

$$\Theta_t = R_{f,t+1}\sqrt{\mathrm{Var}_t\left(M_{t+1}\right)}. \tag{7}$$

This Sharpe ratio can be attained if there are investment strategies with payoffs given by

$$D_{s,t+1} + P_{s,t+1} = c_{0,t} - c_{1,t}M_{t+1}, \quad c_{1,t} > 0, P_{s,t} \geq 0, \tag{8}$$

for scalar $c_{0,t}$ and $c_{1,t}$. The only restriction on $c_{0,t}$ is that it is sufficiently large so that $P_{s,t}$ is nonnegative. One such strategy is a zero-cost portfolio that shorts a single-period instrument that pays off $c_{1,t}M_{t+1}$ and invests the short-sale proceeds in the riskfree asset.

3.2 Sharpe ratios using log excess returns

Sharpe ratios can be defined using log returns for instruments with strictly positive cum-dividend value processes. Denoting the log return to instrument i as $r_{i,t+1}$, the Sharpe ratio using log returns is

$$s_t^i \equiv \frac{E_t r_{i,t+1} - r_{f,t+1} + \frac{1}{2}\mathrm{Var}_t(r_{i,t+1})}{\sqrt{\mathrm{Var}_t(r_{i,t+1})}}. \tag{9}$$

The requirement of a strictly positive cum-dividend process is quite restrictive. It typically rules out portfolios that contain a short position in one or more instruments. Because short positions commonly appear in portfolios designed to maximize Sharpe ratios, simple returns are necessarily used in such a context. For example, Gibbons, Ross, and Shanken (1989) and Hansen and Jagannathan (1991) examine portfolios that maximize the simple-return Sharpe ratio given a vector of expected returns to assets and a covariance matrix of the asset's return shocks.

However, log returns are more tractable than simple returns when studying the properties of individual assets in conditionally Gaussian models. Hence log returns are commonly used in both term structure models and general consumption-based asset pricing models, such as those surveyed by Campbell (2003). Assume that the SDF and asset i's return are both lognormally distributed, so that

$$\begin{pmatrix} \log M_{t+1} \\ \log R_{i,t+1} \end{pmatrix} \equiv \begin{pmatrix} m_{t+1} \\ r_{i,t+1} \end{pmatrix} \sim MVN\left(E_t\begin{pmatrix} m_{t+1} \\ r_{i,t+1} \end{pmatrix}, \mathrm{Var}_t\begin{pmatrix} m_{t+1} \\ r_{i,t+1} \end{pmatrix}\right). \tag{10}$$

Recall that when x_i and x_j are jointly normally distributed with moments

$$E(x_i) = \mu_i, \quad \text{Var}(x_i) = \sigma_i^2, \quad \text{Cov}(x_i, x_j) = \sigma_{ij}, \tag{11}$$

then the corresponding moments for $\exp(x_i)$ and $\exp(x_j)$ are

$$\begin{aligned} E\left(e^{x_i}\right) &= e^{\mu_i + \sigma_i^2/2}, \\ \text{Cov}\left(e^{x_i}, e^{x_j}\right) &= \left(e^{\sigma_{ij}} - 1\right) e^{\mu_i + \mu_j + (\sigma_i^2 + \sigma_j^2)/2}. \end{aligned} \tag{12}$$

Standard algebraic manipulation produces a Sharpe ratio formula for log returns,

$$s_t^i = -\text{Cor}_t(r_{i,t+1}, m_{t+1})\sqrt{\text{Var}(m_{t+1})} \tag{13}$$

Sharpe ratios using simple returns and log returns in (6) and (13) are not equal, although they will be close to each other for expected log returns and variances in the neighborhood of zero.

A lower-case θ denotes the maximum Sharpe ratio using log returns,

$$\theta_t = \sqrt{\text{Var}_t\left(m_{t+1}\right)}. \tag{14}$$

This ratio can be attained if there are investment strategies with payoffs given by

$$\log\left(D_{l,t+1} + P_{l,t+1}\right) = c_{0,t} - c_{1,t} m_{t+1}, \quad c_{1,t} > 0, \tag{15}$$

where $c_{0,t}$ is unrestricted. The payoff is strictly positive, thus $P_{l,t} > 0$. These strategies do not attain the maximum Sharpe ratio defined for simple returns.

3.3 Conditional and unconditional Sharpe ratios

The Sharpe ratios defined in (3) and (9) are conditional Sharpe ratios. More precisely, they are period-t calculations for returns in $t+1$. Empirical analyses often work with unconditional Sharpe ratios, and implicitly treat conditional means and variances as constant. Dynamic term structure models allow for the estimation of conditional Sharpe ratios.

For brevity, whenever this paper refers to a Sharpe ratio, it is a conditional Sharpe ratio unless otherwise noted. Unconditional Sharpe ratios are seldom discussed outside of Section 5's preliminary empirical analysis. The main use of the term "conditional" is when referring to the unconditional properties of conditional Sharpe ratios. An example that plays an important role in the empirical analysis is the sample mean of (14), which is sample unconditional mean of the conditional maximum Sharpe ratio using log returns. This sample mean should not be confused with the sample unconditional Sharpe ratios.

4 Sharpe ratios in Gaussian term structure models

This section describes the Gaussian term structure framework. It is designed to set up notation and formulas. Nothing in it is original, therefore the model is only briefly sketched in Section 4.1. Formulas for Sharpe ratios of bonds and bond portfolios are presented in Section 4.2, and formulas for maximum Sharpe ratios are presented in Section 4.3.

4.1 A Gaussian model

The starting point of most term structure models is a state vector that determines the dynamics of the SDF. There is a length-n state vector x_t that follows a Gaussian vector autoregression. The dynamics of the state are

$$x_{t+1} = \mu + K x_t + \Sigma \epsilon_{t+1}, \quad \epsilon_{t+1} \sim MVN(0, I). \tag{16}$$

The continuously-compounded riskfree rate is an affine function of the state vector,

$$r_t = \delta_0 + \delta_1' x_t. \tag{17}$$

The log SDF has the form

$$m_{t+1} = -r_t - \frac{1}{2}\Lambda_t'\Lambda_t - \Lambda_t'\epsilon_{t+1}. \tag{18}$$

Thus the SDF is conditionally lognormally distributed with conditional mean and variance

$$E_t m_{t+1} = -r_t - \frac{1}{2}\Lambda_t'\Lambda_t, \qquad \text{Var}_t\left(m_{t+1}\right) = \Lambda_t'\Lambda_t. \tag{19}$$

The vector Λ_t is the period-t compensation investors require to face factor risk.

To compute bond prices, we must specify the functional form of Λ_t. The essentially affine version, introduced in Duffee (2002), is

$$\Lambda_t = \Sigma^{-1}\left(\lambda_0 + \lambda_1 x_t\right). \tag{20}$$

The essentially affine equivalent-martingale dynamics of x_t are

$$x_{t+1} = \mu^q + K^q x_t + \Sigma \epsilon^q_{t+1}, \quad \epsilon^q_{t+1} \sim MVN(0, I), \tag{21}$$

where

$$\mu^q = \mu - \lambda_0, \quad K^q = K - \lambda_1. \tag{22}$$

Denote the price, log price, and continuously-compounded yield on a k-maturity zero-coupon bond by $P_t^{(k)}, p_t^{(k)}$, and $y_t^{(k)}$ respectively. Applying the intuition of Duffie and Kan (1996), Ang and Piazzesi (2003) show that log bond prices in this setting are affine in the state vector. Write the log bond price as

$$p_t^{(k)} = A_k + B_k' x_t. \tag{23}$$

The loading of the log price on the state vector is

$$B_k' = -\delta_1' \left(I - K^q\right)^{-1} \left(I - \left(K^q\right)^k\right) \tag{24}$$

and the constant term satisfies the difference equation

$$A_1 = -\delta_0, \quad A_{k+1} = -\delta_0 + A_k + B_k' \mu^q + \frac{1}{2} B_k' \Sigma \Sigma' B_k. \tag{25}$$

4.2 Sharpe ratios for bonds and bond portfolios

Log returns to individual bonds are easy to analyze in this model. The log return to a k-period bond from t to $t+1$ is normally distributed and given by

$$p_{t+1}^{(k-1)} - p_t^{(k)} = r_t + B_{k-1}'(\lambda_0 + \lambda_1 x_t) - \frac{1}{2} B_{k-1}' \Sigma \Sigma' B_{k-1} + B_{k-1}' \Sigma \epsilon_{t+1}. \tag{26}$$

Denote the time-invariant standard deviation by

$$\sigma^{(k)} \equiv \left(B_{k-1}' \Sigma \Sigma' B_{k-1}\right)^{1/2}. \tag{27}$$

For future reference, denote this log return's conditional mean and covariance with the return to a j-period bond by

$$\mu_t^{(k)} \equiv r_t + B_{k-1}'(\lambda_0 + \lambda_1 x_t) - \frac{1}{2} \left(\sigma^{(k)}\right)^2, \qquad \sigma^{(k,j)} \equiv B_{k-1}' \Sigma \Sigma' B_{j-1}. \tag{28}$$

The bond's Sharpe ratio using log returns is

$$s_t^{(k)} = \frac{B'_{k-1}(\lambda_0 + \lambda_1 x_t)}{\sigma^{(k)}}. \tag{29}$$

Because this Sharpe ratio is linear in x_t, its unconditional properties are analytically tractable. For example, the unconditional mean of (29) simply replaces the realization of the state in the numerator with its unconditional expectation. Note that the unconditional mean of (29) is the unconditional expectation of the bond's conditional Sharpe ratio, not the bond's unconditional Sharpe ratio. The unconditional Sharpe ratio using log returns is

$$s^{(k)} = \frac{B'_{k-1}(\lambda_0 + \lambda_1 E(x_t))}{\left((\sigma^{(k)})^2 + B'_{k-1}\lambda_1 \mathrm{Var}(x_t)\lambda'_1 B_{k-1}\right)^{1/2}}. \tag{30}$$

The two terms in the denominator of (30) are the one-period-ahead conditional variance of the bond's excess return and the unconditional variance of the one-period-ahead conditional mean excess return.

Sharpe ratios of bond portfolios must be studied using simple returns, because a portfolio of assets with lognormally-distributed returns does not have a lognormally-distributed return. For a given factor realization x_t, Sharpe ratios using simple returns can be calculated easily for both individual bonds and bond portfolios.

Denote the simple one-period return to a k-period bond by $R_{t+1}^{(k)}$. Using properties of the lognormal distribution in (12), the period-t expectation of the simple excess return is

$$E_t\left(R_{t+1}^{(k)}\right) - R_{f,t+1} = \exp\left(\mu_t^{(k)} + \frac{1}{2}\left(\sigma^{(k)}\right)^2\right) - \exp\left(r_t\right). \tag{31}$$

The conditional covariance of this return with the simple excess return to a j-period bond is

$$\mathrm{Cov}_t\left(R_{t+1}^{(k)}, R_{t+1}^{(j)}\right) = \left(\exp\left(\sigma^{(k,j)}\right) - 1\right)\exp\left(\mu_t^{(k)} + \mu_t^{(j)} + \frac{1}{2}\left(\left(\sigma^{(k)}\right)^2 + \left(\sigma^{(j)}\right)^2\right)\right). \tag{32}$$

Denote the excess return vector to d zero-coupon bonds by R_{t+1}^e. Denote the conditional covariance matrix of the excess returns by Ω_t. Although the dimension of the state vector is n, the covariance matrix of simple returns can have rank greater than n because bond prices are nonlinear functions of the state. Denote the dollar investments in each bond by a vector ρ_t. The Sharpe ratio, using simple returns, of the portfolio is

$$S_t^p = \frac{\rho'_t E_t(R_{t+1}^e)}{\left(\rho'_t \Omega_t \rho_t\right)^{1/2}}. \tag{33}$$

4.3 Maximum Sharpe ratios

Straightforward algebra reveals that for this Gaussian model, the maximum Sharpe ratio based on simple excess returns is

$$\Theta_t = \sqrt{\exp\left(\Lambda_t' \Lambda_t\right) - 1}. \tag{34}$$

Similar calculations produce the maximum Sharpe ratio based on log returns,

$$\theta_t = \sqrt{\Lambda_t' \Lambda_t}. \tag{35}$$

Thus the squared maximum Sharpe ratio using log returns is a first-order Taylor series approximation to the squared ratio using simple returns. The expansion is around the point $\text{Var}_t(m_{t+1}) = 0$.

Investments that attain the former Sharpe ratio have payoffs

$$D_{s,t+1} + P_{s,t+1} = c_{0,t} - c_{1,t} \exp\left(-\Lambda_t' \epsilon_{t+1}\right), \quad c_{1,t} > 0, P_{s,t} \geq 0. \tag{36}$$

Investment strategies that attain the latter Sharpe ratio have payoffs

$$\log\left(D_{l,t+1} + P_{l,t+1}\right) = c_{0,t} + c_{1,t} \Lambda_t' \epsilon_{t+1}, \quad c_{1,t} > 0. \tag{37}$$

In general, investment strategies in bond portfolios cannot attain either maximum Sharpe ratio.[2] However, if the financial market allows the trading of instruments with payoffs that are arbitrary functions of bond prices, then outside of special cases such fixed-income instruments can attain both maximum Sharpe ratios.

The payoffs of the derivative instruments are written in terms of log bond prices. Stack log prices of n bonds into the vector p_t. (Recall that n is the dimension of the state vector.) Using the notation

$$p_t = A + B x_t, \tag{38}$$

where A and B are stacked versions of A_k and B_k in (23), spanning requires that B is invertible. For general Gaussian models this requirement is satisfied, although Duffee (2008) notes that in special cases B is singular. One fixed-income instrument that attains the maximum period-t conditional Sharpe ratio for simple returns has a single payoff at $t + 1$

[2] This is obvious for log returns, since bond portfolios do not have normally-distributed log returns. Section 6 discusses why maximum Sharpe ratios produced using bond portfolios can be much less than maximum Sharpe ratios implied by an SDF.

given by

$$D^B_{s,t+1} = c_{0,t} - \exp\left(-\Lambda_t' \Sigma^{-1} B^{-1}(p_{t+1} - A)\right). \tag{39}$$

A fixed-income instrument that attains the maximum Sharpe ratio for log returns can be constructed in similar fashion.

Standard portfolio mathematics tells us that for simple returns, the maximum Sharpe ratio for the bond portfolio in (33) is

$$\Theta^p_t = \sqrt{E_t(R^e_{t+1})' \Omega_t^{-1} E_t(R^e_{t+1})}. \tag{40}$$

Similarly, the unconditional maximum Sharpe ratio for a fixed-weight portfolio is

$$\Theta^p = \sqrt{E(R^e)' \mathrm{Var}(R^e)^{-1} E(R^e)}. \tag{41}$$

Monte Carlo simulations can be used to study the properties of these Sharpe ratios.

5 A preliminary look at unconditional Sharpe ratios

Two broad themes are apparent in the literature analyzing unconditional returns to Treasury securities. First, unconditional Sharpe ratios of long-term Treasury securities are low relative to unconditional Sharpe ratios of equity portfolios. A classic reference is Fama and French (1993), who estimate a monthly unconditional Sharpe ratio of 0.02 for bonds with maturities from six to ten years. They refer to the risk premium as "puny." Second, unconditional Sharpe ratios of short-maturity Treasury bills are suspiciously high; see, e.g., Hansen and Jagannathan (1991). High excess returns at the very short end are typically attributed to bid-ask spreads as in Luttmer (1996) or idiosyncratic market conditions as in Duffee (1996).

Undoubtedly part of the computed Sharpe ratios at maturities near zero are illusory, in the sense that they overstate Sharpe ratios that large investors can attain by trading in those instruments. (I drop the adjective "unconditional" because every Sharpe ratio in this section is unconditional.) But we should not be quick to conclude that Sharpe ratios of longer-term instruments are a good measure of attainable Sharpe ratios on shorter-term instruments. Even if we ignore the shortest-maturity instruments, Sharpe ratios are inversely related to maturity. For example, we can infer from other estimates in Fama and French that a portfolio containing Treasury bonds with less than five years to maturity has a monthly Sharpe ratio that slightly exceeds the monthly Sharpe ratio of the stock market. Evidence in Campbell and Viceira (2001) also supports the inverse relation between Sharpe ratios and maturity.[3]

[3]Notwithstanding its title, Campbell and Viceira (2005) does not discuss Sharpe ratios across maturities.

But the largest Sharpe ratio they calculate using quarterly data from 1952 to 1996 is not much more than half the Sharpe ratio for the stock market.

Table 1 updates and extends the relevant evidence of Fama/French and Campbell/Viceira. Treasury bonds are sorted by maturity and placed in seven portfolios. The short end is a one-month to six-month bucket and the long end is a five-year to ten-year bucket. The data are from the Center for Research in Security Prices (CRSP) and span the period January 1952 through December 2008. Following Fama and French, the table reports means and standard deviations of simple excess monthly returns, along with corresponding Sharpe ratios. The riskfree rate is the return to one-month Treasury bills, as calculated by Ibbotson Associates and made available on the website of Ken French. Following Campbell/Viceira, the table also reports these values at a quarterly frequency.[4] The riskfree rate is the return to three-month Treasury bills from CRSP.

The table supports three main conclusions. First, it confirms that the inverse relation between maturity and Sharpe ratio holds across the term structure. Second, after adjusting for the horizon (divide the quarterly values by $\sqrt{3}$), reported Sharpe ratios at the monthly horizon are higher than those at the quarterly horizon. Third, Sharpe ratios for bonds with maturities less than two years are close to or exceed the stock market's Sharpe ratio at both monthly and quarterly horizons. At the quarterly horizon, this last result differs from Campbell and Viceira because of an additional 12 years of returns that were kinder to bondholders than stockholders.

The high Sharpe ratios at the monthly horizon relative to those at the quarterly horizon are driven by the use of the one-month bill yield. Effectively, the monthly Sharpe ratios are calculated assuming the one-month yield is a rate at which investors can borrow and lend risklessly. Given the evidence of Duffee (1996), a reasonable view is that owing to market imperfections—in particular, the inability of investors to issue their own bills—the yield on this bill is typically lower than the rate at which market participants can borrow and lend risklessly. Hence after adjusting for the pure time effect of the horizon, true monthly Sharpe ratios on individual bonds are probably better calculated from the quarterly results. Therefore the next section uses maximum unconditional monthly Sharpe ratios between 0.15 and 0.20 as benchmarks for a complete bond market.

Sharpe ratios of the stock market are not studied further in this paper. But it is worth emphasizing that the maximum Sharpe ratio of a portfolio of Treasury bonds comfortably exceeds the Sharpe ratio for the market. Nonetheless, the book "Treasury securities for the

Its use of the phrase "term structure" refers to investors with heterogeneous investment horizons.

[4]The table uses simple returns, while Campbell and Viceira use log returns. Quarterly returns are built up from rolling over monthly positions in the portfolios.

long run" is unlikely to be a bestseller. To match mean excess returns to stocks, investors will have to leverage up to buy portfolios of Treasury bonds. Following Luttmer (1996), the trading costs of frequently rolling over these positions will erode the high Sharpe ratios for all but large financial institutions.

6 Empirical analysis

This long section describes the results of estimating Gaussian term structure models. The number of factors ranges from two to five. The models are estimated with maximum likelihood, with or without a constraint on the model's maximum Sharpe ratio. The first two subsections describe the data and estimation technique for unconstrained models. The next two subsections summarize some important properties of the estimated unconstrained models. Section 6.5 describes how the models are estimated subject to a Sharpe ratio constraint and Sections 6.6 and 6.7 summarize properties of the estimated constrained models. Section 6.8 interprets the results in terms of principal components of yields, while Section 6.9 compares forecast accuracy of many of the estimated models.

6.1 Data for term structure estimation

The empirical implementation treats each period as a month. The models are estimated using a monthly panel of eight zero-coupon Treasury bond yields. Because of the problems with one-month yields mentioned in Section 5, the shortest-maturity yield I use is the three-month yield (bid/ask average) from CRSP. Artificially-constructed yields on zero-coupon bonds with maturities of one, two, three, four, and five years are also from CRSP. Yields on six-year and ten-year bonds are from the Federal Reserve Board's website and are constructed using the procedure of Gurkaynak, Sack, and Wright (2006). The sample period is November 1971 through December 2008. The starting month is the month when the cross-section of Federal Reserve data expands.

6.2 Estimation methodology for unconstrained models

Estimation and hypothesis testing are performed with maximum likelihood via the Kalman filter. Stack the observed yields in the d-vector y_t, which are observed in months $1, \ldots, T$. Using state-space language, the transition equation of the underlying state is (16) and the measurement equation is

$$y_t = A + Bx_t + \eta_t, \quad \eta_t \sim MVN(0, \sigma_\eta^2 I). \tag{42}$$

In (42), A is a d-vector and B is a $d \times n$ matrix. They are determined by the Duffie-Kan restrictons (24) and (25).

The transition and measurement equations are underidentified because the state vector is latent. For identification, the vector can be arbitrarily scaled, rotated, and translated. For estimation purposes, I normalize the transition equation (16). The constant term μ is zero, the feedback matrix K is diagonal, and the lower-triangular volatility matrix Σ has ones along the diagonal. Thus there are $n+n(n-1)/2$ free parameters in the transition equation, $n+1$ free parameters in the short-rate equation (17), $n+n^2$ additional free parameters in the equivalent-martingale dynamics (21), and a final free parameter σ_η in the measurement equation (42).

The unconstrained n-factor model is estimated for $n = 2, \ldots, 5$. The number of free parameters ranges from 13 for $n = 2$ to 52 for $n = 5$. The estimated values of the individual parameters are not of direct interest here, and thus are not reported. Instead, I focus on features of the estimated models related to Sharpe ratios.

6.3 Sharpe ratios of unconstrained models

This subsection summarizes properties of four different Sharpe ratios, as implied by the estimated models. Two are the maximum conditional Sharpe ratios attainable with a complete fixed-income market. The simple return version is (34) and the log return version is (35). The other two are maximum conditional and maximum unconditional Sharpe ratios attainable with a complete zero-coupon bond market for maturities up to ten years.

This hypothetical bond market contains 120 bonds with maturities ranging from one to 120 months. Not all of these bonds are needed to form a complete bond market, at least up to machine precision. (Recall that simple returns are nonlinear functions of the state, thus more than n bonds are needed.) More precisely, the covariance matrices of returns in (40) and (41) are singular, up to machine precision, with such a large portfolio. Trial and error is used to determine a set of bonds, for each n, that spanned the bond market for maturities up to ten years.

I calculate both sample and population means of the three conditional maximum Sharpe ratios discussed here. Sample means use filtered values of the state vector. For each estimated model, the Kalman filter produces a time series of filtered values, denoted $\hat{x}_t$. The filtered state, combined with parameter estimates, implies a filtered estimate $\hat{\Lambda}_t$ from (20). If we assume that the fixed-income market is complete, there is an admissible investment strategy that achieves the Sharpe ratio for simple returns $\hat{\Theta}_t$ from (34) and for log returns $\hat{\theta}_t$ from (35). Similarly, given the filtered state and parameter estimates, conditional mean excess simple bond returns and the conditional covariance matrix among these returns are produced following the steps in Section 4.2 and maximum Sharpe ratios $\hat{S}_t^p$ are then calculated using (40).

Population means are calculated using Monte Carlo simulations. For each estimated model, a simulated time series of 300,000 months is generated and the population properties of the conditional Sharpe ratios are computed. The simulation is also used to compute the maximum unconditional Sharpe ratio of the complete bond market.

Panel A of Table 2 reports the relevant information. The table is easy to summarize. The Sharpe ratios for the two-factor model are a little low. The ratios for the three-factor model are a little high, and those for the four-factor and five-factor models are impossibly large. Recall that Section 5 set a benchmark between 0.15 and 0.20 for the unconditional maximum Sharpe ratio of the bond market. The population value for the three-factor model is 0.19, while the values for the four-factor and five-factor models are around two to three times the benchmark range, respectively. The plausibility of these high-dimensional models drops further when we look at means of maximum conditional Sharpe ratios.

Assuming a complete fixed-income market, the sample mean of the maximum conditional Sharpe ratio is about 0.8 for the four-factor model and about 10^{30} for the five-factor model. Since the former value looks quite modest by comparison, it is helpful to recall it corresponds to an annual Sharpe ratio of 2.7.

What is the source of the bizarre five-factor, complete-market Sharpe ratio? In a nutshell, it is the nonlinearity inherent in the SDF. Because the SDF is bounded below by zero, high SDF volatility corresponds to a highly skewed SDF. Simple monthly returns to bonds are close to linear in the state vector. Thus the greater the skewness, the lower the conditional correlation between the SDF and the return to a portfolio of bonds. The appendix discusses this point in detail in the context of a one-factor model.

The four-factor and five-factor models imply substantial predictability of excess bond returns, which in turn implies quite high conditional maximum Sharpe ratios for bond portfolios. Table 2 reports the sample means of these monthly Sharpe ratios for the four-factor and five-factor models are close to 0.6 and 2.8 respectively, or the annual equivalent of 2.0 and 9.7. For the five-factor model, the largest value in the sample of 446 months is about 48. The only way the model can reproduce this bond-market Sharpe ratio is with an SDF that is astronomically volatile, and has near-zero correlations with returns to bond portfolios.

Nonlinearities disappear when working with log returns and the log SDF. Hence, as Table 2 reports, mean maximum Sharpe ratios using log returns are not eye-popping for the five-factor model, although they remain totally unrealistic. It is worth noting, though, that the information in the maximum Sharpe ratios for log returns is identical to that in the maximum

Sharpe ratios for simple returns; they are monontonic transformations of each other.

Figure 1 displays the time series of filtered maximum Sharpe ratios using log returns, assuming a complete fixed-income market.[5] The models disagree substantially about the periods when Sharpe ratios are high and when they are low. The correlation between the two-factor estimates and the three-factor estimates is 0.72, but the correlation between the two-factor and five-factor estimates is only 0.03.

A natural concern of the reader is that the absurdly high Sharpe ratios of the five-factor model are somehow an artifact of this paper's data sample or estimation procedure. Perhaps Sharpe ratios would be lower if the three-month bill yield is excluded, restricting the analysis to longer-maturity yields. Alternatively, perhaps shifting the focus to annual changes in yields rather than monthly changes produces more sensible results. Conveniently, the on-line appendix to Cochrane and Piazzesi (2005) describes how to construct analytically a no-arbitrage dynamic model that replicates the results from their annual-horizon vector autoregression. The appendix to this paper follows their procedure, which uses bonds with maturities of one through five years. The resulting implied time series of maximum Sharpe ratios using simple returns has a sample mean of about 10^{18}.

An obvious conclusion is that unconstrained term structure models with more than three factors wildly overfit the data. One approach to this problem is to impose parameter restrictions prior to estimation. For example, Cochrane and Piazzesi (2008) construct a four-factor model in which there is a single priced risk and a single factor that drives variation in that priced risk. Joslin, Priebsch, and Singleton (2009) construct a five-factor model with a variety of parametric restrictions on risk compensation. If we are confident that we understand how risk premia vary over time, this approach is probably the best way to avoid overfitting. But because I am unwillng to make *a priori* assumptions about the precise dynamic behavior of risk premia, I instead impose a Sharpe ratio constraint. To discuss how the constraint affects the results of term structure estimation, we first need to look at a few more properties of unconstrained models.

6.4 Other features of unconstrained models

Although the four estimated unconstrained models differ substantially in their Sharpe ratio implications, they agree on many unconditional properties of the term structure over the range of maturities used to estimate the model. For maturities from three months to ten years, the models generate roughly matching mean yields, mean standard deviations of yields, and mean conditional Sharpe ratios. They also agree on the persistence of long-maturity

[5]Simple returns are not used because the scale of the figure for the five-factor model is meaningless.

yields. The two main areas of disagreement are (a) means and standard deviations of yields outside of this range of maturities—in particular, at the short end of the yield curve; and (b) the fraction of variability in yields that is forecastable. These disagreements drive the Sharpe ratio patterns across the estimated models.

Surprisingly (at least at first glance), the models agree that unconditional mean yields from three months to ten years are about one percentage point per year *lower* than the corresponding means in the data sample. The relevant evidence is in Table 3. The last row of the table reports sample mean yields, while the first four rows report unconditional mean yields implied by the estimated unconstrained models. (For now, ignore the rows containing results for constrained models.) For example, the ten-year yield has a sample mean of about 7.5 percent per year and population means of about 6.5 percent across all four estimated models. The reason for this discrepancy is that on average, yields fell during the sample period. The term structure was about four percentage points higher in November 1971 than in December 2008. The autoregressive data-generating process interprets this decline as yields reverting toward their means, thus unconditional means are less than sample means.

The relevant evidence about standard deviations is also in Table 3. For example, model-implied one-month-ahead standard deviations of shocks to the ten-year yield range from 0.32 to 0.33 annual percentage points across the four unconstrained models. Unconditional standard deviations of the same yield range from 2.6 to 2.7 percentage points. The sample standard deviation is 2.4 percentage points. This discrepancy is typical with highly persistent data. An estimate of this persistence is in Table 4. The table reports that model-implied unconditional correlations between the month-t ten-year yield and the month $t+120$ ten-year yield is about 0.3 across the four models.

As Section 6.6 discusses, a major difference between unconstrained and constrained estimates is the extent to which investors could foresee the drop in bond yields during the 1990s. To set the stage, consider December 1988, at the end of the Reagan administration. The ten-year yield was 9.0 percent. According to all of the unconstrained models, investors at that time expected a yield of 7.1 percent twelve years later. (This information is not reported in any table.) The actual yield in December 2000 was 5.3 percent. Put differently, half of the actual decline was predicted by investors.

The one-month yield is outside of the range of maturities in estimation. Its implied behavior depends critically on the number of factors in the model. Table 3 reports that the mean slope between the one-month and three-month yields rises with n. For the two-factor model, the slope is seven basis points. It rises to 10, 18, and an eye-popping 63 basis points for $n = 3, n = 4$, and $n = 5$ respectively. The yield's one-month-ahead standard deviation ranges from 0.6 to 1.3 annual percentage points across the unconstrained models.

The mean slope of the yield curve affects mean Sharpe ratios. Investors who buy a three-month bond and hold it for a month expect to profit by sliding down the yield curve. Table 4 reports model-implied unconditional mean conditional Sharpe ratios for three-month, five-year, and ten-year bonds. (Log returns are used.) For the five-year and ten-year bonds, the models roughly agree. Estimates for the former range from 0.09 to 0.11, while estimates for the latter range from 0.07 to 0.09. But for the three-month bond, the estimates range from 0.11 with two factors to an astronomical 0.75 with five factors.

The other main difference across the unconstrained models is the amount of predictability in yields (and returns). A detailed discussion of this predictability is deferred to Section 6.7. Here it is sufficient to note that the unconditional standard deviations of bonds' conditional Sharpe ratios rise substantially with the number of factors. For example, Table 4 reports that the standard deviation of the ten-year bond's conditional Sharpe ratio ranges from 0.15 for the two-factor model to 0.31 for the five-factor model. The next subsections describe how this predictability depends on the constraint imposed on maximum Sharpe ratios.

6.5 Constrained model estimation

I estimate the models with maximum likelihood subject to the constraint

$$\overline{\hat{\theta}_t} \leq c. \tag{43}$$

In words, the sample mean of the filtered conditional maximum Sharpe ratios (using log returns) cannot exceed the scalar c. Alternatively (or additionally), constraints could be placed on population means or on maximum Sharpe ratios computed with simple returns. The choice of (43) is dicated by its coding simplicity.

For a given number of factors n, the constraint c steps down in increments of 0.015 from the unconstrained sample mean reported in Table 2 to no less than 0.05. For example, for $n = 4$, I estimate the model 33 additional times, each time tightening the constraint by 0.015 until it reaches $c = 0.061$. In practice, a fairly complicated algorithm is employed to ensure that the numerical optimization procedure locates the global maximum.

Some information from the sequences of parameter estimates for $n = 4$ and $n = 5$ are displayed in Figures 2 and 3. For example, Panel A of Figure 2 displays the effect of the constraint on the log-likelihood. Recall that for a five-factor model, the unconstrained sample mean of $\hat{\theta}_t$ is about 1.8. The figure shows that effect of the constraint on the five-factor log-likelihood is negligible for $c > 0.8$. At its tightest, the constraint of $c = 0.05$ produces a log-likelihood 70 below the unconstrained log-likelihood.

In addition to these sequences of constrained estimations, the model is estimated subject to $c = 0.25$ for each of $n = 3, 4$, and 5. Detailed information about these estimated models is presented in various tables, including Panel B of Table 2. Trial and error revealed that this choice produces unconditional maximum Sharpe ratios for a complete bond market that are close to the benchmark range, from Section 5, of 0.15 to 0.20. The table reports that the population value of this ratio is 0.18 for the constrained three-factor and four-factor models and 0.17 for the constrained five-factor model. Additional features of these models are discussed in the following subsections.

6.6 Sharpe ratios and the unconditional term structure

The level and shape of the unconditional mean yield curve are sensitive to the Sharpe ratio constraint. Tightening the constraint raises mean yields for all maturities and flattens the slope at the short end. Some visual evidence is in Panel B of Figure 2, which displays the unconditional mean short rate as a function of the tightness of the constraint c. Tightening the constraint raises the mean short rate. At the tightest value of c, the mean short rates are both above 8.5 percent. Means of other yields are reported in Table 3 for constrained models that impose $c = 0.25$. The wedge between unconstrained and constrained models is largest for the five-factor model. For the constrained model, mean three-month and ten-year yields are 9.3 and 10.3 percent respectively. The higher mean yields reduce mean conditional Sharpe ratios because they effectively convert expected declines in interest rates (and thus expected capital gains on long-maturity bonds) into positive shocks to capital gains.

Reconsider investor expectations, as of December 1988, of the ten-year yield in December 2000. The unconstrained models imply investors predict about a two percentage point drop in yields. Imposing $c = 0.25$ wipes this out. For the three-factor model, where the constraint is least binding, investors anticipate only a 50 basis point drop in yields. For the four-factor and five-factor models, investors actually anticipate an *increase* in yields of between 50 and 90 basis points, respectively.

Constraining Sharpe ratios also lowers the spread between mean three-month and one-month yields, thus lowering mean Sharpe ratios for short-maturity bonds. Some visual evidence is in Figure 2. Panel C reports the unconditional mean Sharpe ratio for a three-month bond. It declines almost linearly in c; the slope is about 0.5. Panel D illustrates that the constraint has a much smaller effect on the mean unconditional Sharpe ratio of a ten-year bond. Again the relation is approximately linear, but the slope is closer to 0.05. Similar evidence is reported in Table 3 for $c = 0.25$. For each value of n, the mean Sharpe ratio of the three-month bond is about 0.16 and the mean Sharpe ratio of the ten-year bond is about 0.06.

Note that in Figure 2, the mean Sharpe ratio of the short-maturity bond exceeds that of the long-maturity bond for each value of c. Put differently, the conclusion that mean Sharpe ratios are inversely related to maturity is robust to beliefs about about the maximum plausible Sharpe ratio. We return to this result in Section 6.8.

6.7 Return predictability and time-variation in Sharpe ratios

In Gaussian models, Sharpe ratios vary over time because of predictability in excess returns. The four-factor and five-factor models estimated here exhibit three basic types of excess return predictability. The first is short-horizon predictability in short-maturity bond returns. The second is short-horizon predictability in long-maturity bond returns. The third is long-horizon predictability in long-maturity bond returns. I discuss them in turn.

Short-horizon return predictability to short-maturity bonds is probably the most dramatic but economically the least interesting. As discussed in Section 6.4, the high-dimensional unconstrained models imply unrealistic properties of very short-maturity yields, which are not used in estimating the model. One of these unrealistic properties is that the forward rate from two months to three months swings wildly from month to month, generating large swings in predicted excess returns and thus large variations in Sharpe ratios. Panel A of Figure 3 reports population R^2s of regressions of monthly excess log returns to a three-month bond on the lagged factors. Unless Sharpe ratios are tightly constrained, the R^2s for the four-factor and five-factor models exceed 10 percent. These high R^2s are equivalent to the large standard deviations of conditional Sharpe ratios for the three-month bond, as discussed in Section 6.4. As reported in Table 4, the constraint $c = 0.25$ damps considerably both the mean and the standard deviation of this conditional Sharpe ratio. For the constrained models, the mean ranges from 0.15 to 0.17, and the standard deviation ranges from 0.15 to 0.22.

Over the range of maturities used in estimation, the unconstrained models generate much more sensible (i.e., smaller) excess return predictability. Table 5 reports population R^2s of regressions of monthly excess returns to a ten-year bond on the lagged factors. They range from 2.2 percent for the two-factor model to 8.9 percent for the unconstrained five-factor model. Tightening the constraint has a relatively small effect, as displayed in Panel B of Figure 3. Table 3 also reports that when the constraint is set to $c = 0.25$, the models all generate population R^2s in the neighborhood of three to four percent.

How many factors drive expected excess returns? To begin answering this question, I follow Joslin et al. (2009) and construct model-implied principal components of the covariance matrix of conditional expected excess monthly returns to a set of bonds. I use bonds with maturities of 2, 3, ..., 10 years. The first principal component explains more than 99.5 percent of the variation in these monthly excess returns. This result holds regardless of the number of factors and regardless of the constraint imposed on Sharpe ratios. (Panel C of Figure 3 plots this percentage for four-factor and five-factor models, but the figure simply looks like it has a rather thick line at the point where the fraction equals one.) Thus a robust conclusion is that monthly excess returns are driven by a single factor (outside of very short-maturity bonds).

But when we turn to annual excess returns, the results are more complicated. Aside from the two-factor model, the unconstrained models all agree that about 20 percent of annual excess returns to a ten-year bond are predictable. The evidence is in Table 5. Constraining Sharpe ratios modestly reduces this predictability, again as shown in Table 5. With $c = 0.25$, the population R^2s are in the range of 15 to 19 percent. The complication shows up in the fraction of predictable annual excess returns that are explained by the first principal component of monthly excess returns. The fraction is reported in the final column of Table 5.

The four-factor and five-factor models imply that a substantial fraction of predictable annual excess returns are orthogonal to the first principal component of monthly excess returns. For the unconstrained four-factor and five-factor models, the amounts are 30 and 50 percent respectively. When $c = 0.25$, the amounts fall to 22 percent and 36 percent respectively. Even for the three-factor model, 15 percent of the predictable variation in annual excess returns is orthogonal to the factor driving predictable monthly returns.

How should we interpret the different kinds of variation in excess returns to long-maturity bonds? To put some structure on this variation, I use a standard decomposition of the term structure.

6.8 Principal components decompositions of yields

Term structure factors can be rotated into level, slope, curvature, and everything else. The perceived importance of the final category has grown substantially since the work of Cochrane and Piazzesi (2005). Here, I use this sort of rotation to interpret the sources of average risk compensation and time-varying risk compensation.

There are three main conclusions. First, the inverse relation between maturity and mean Sharpe ratios is driven by the average compensation investors require to face slope risk. Long-maturity bonds hedge this risk, while short-maturity bonds face this risk. Second, variations in slope of the term structure are associated with fairly long-lived variations in expected excess returns. Finally, the fourth principal component is associated with shortlived variations in these expectations. The factor has a minimal effect on current yields, but significant effects on short-run expectations of current risk premia and future yields.

I document these conclusions by focusing on a single estimated model: the four-factor model that is constrained by $c = 0.25$. Results for the unconstrained four-factor model, as well as constrained and unconstrained five-factor models, are very similar. The four factors are rotated into the four principal components of the unconditional covariance matrix of bond yields.[6]. These factors have the usual properties. The first row of Table 6 reports that the first factor (level) explains close to 98 percent of the total unconditional variation in the term structure. The second factor (slope) picks up two percent, the third factor (curvature) picks up 0.1 percent, and the fourth factor explains virtually none of the unconditional variation. As we will see, it picks up conditional variation in yields.

The shock to the log return on a k-maturity bond is the vector of factor loadings B_{k-1} times the factor shocks. The factors are normalized to mean zero, thus the mean compensation to face shock i is, from (26), element i of B_{k-1} times element i of λ_0. The next set of rows of Table 6 report this product for each factor and three different bonds. The products are divided by the unconditional standard deviation of the bond's log excess return. Summing across the factors produces the mean Sharpe ratio for the bond.

These components of mean Sharpe ratios are easy to interpret. Investors require compensation to face the risk that the level of the term structure unexpectedly rises. The first column tells us that investors receive roughly the same compensation for this risk across the reported range of maturities. Investors also require compensation to face the risk that the slope of the term structure unexpectedly tilts down. Such a tilt lowers the prices of short-maturity bonds (raises their yields) and raises the prices of long-maturity bonds (lowers their yields). Thus short-maturity bonds exacerbate this risk and long-maturity bonds hedge this risk. This pattern explains the cross-sectional variation in risk compensation for the second factor. Mean compensations for the remaining two factors are negligible.

The next set of rows in Table 6 show that the slope and as-yet unnamed fourth factor each account for about half the variation in conditional expected log excess monthly returns. The final set of rows show that the slope factor is much more important in accounting for variations of conditional expected log excess annual returns. It picks up about 85 percent of the total variation. Figure 4 helps to explain the difference between monthy and annual return horizons.

Panels A and B in the figure display impulse responses to the second factor. The factor is assumed to increase by one standard deviation at month zero, holding the other factors

[6]The included yields are those on all maturities from one month to 120 months. A pseudo-inverse is used to compute the principal components.

constant. In Panel A, the month-zero response to the second factor is for short-maturity yields to fall by more than 100 basis points and for long-maturity yields to rise by about 50 basis points. This is shown by the black dashed line. The panel also shows the impulse response for month 12, which is a blue dotted-dashed line. Both short-maturity and long-maturity yields have begun to drift back, but the slope remains steep.

Panel B displays the effect on expected excess monthly returns across the term structure. As is well-known, a steeper slope predicts higher excess returns. For the ten-year bond, the immediate increase in the monthly excess return is close to 50 basis points. Twelve months later, expected excess returns remain 35 basis points above normal.

Similar information is displayed in Panels C and D for the fourth factor. Panel C shows that the factor has a trivial month-zero effect on yields. At most, there are some wiggles at the short end. However, the factor predicts that during the next twelve months, the level of the term structure will drop by around 20 to 25 basis points. Since month-zero yields are unchanged, month-zero expected excess returns accordingly jump. For the ten-year bond, the increase in expected excess return is about 40 basis points. Twelve months later, expected excess returns are back to normal.

The results summarized here, as well as in Section 6.7, complicate the econometrician's job of term structure estimation. It is straightforward to impose parametric assumptions on a model, such as those adopted by Cochrane and Piazzesi (2008), to limit the model's flexibility. They assume a single priced risk and a single factor that creates variation in the price. But the evidence here points to multiple priced risks and multiple sources of variation in these prices.[7] Similarly, the results of Section 6.6 illustrate an important shortcoming in the common assumption, made for estimation convenience, that factors' unconditional means equal their sample means. Such a restriction has hidden, but important and perhaps unwanted implications for Sharpe ratios.

Thus these modeling choices affect the way we interpret economically the risk premia of fixed-income instruments. However, from another economic perspective, differences among all of these models—whether a two-factor model or a five-factor model, whether Sharpe ratios are constrained or unconstrained—are barely worth mentioning. This perspective is considered next.

[7]For the four-factor and five-factor models estimated in this paper, the effect of the slope and fourth factors on time-varying risk premia is almost entirely a result of varying the price of level risk. These results are not reported in any table or figure.

6.9 Cross-sectional and forecast errors

How accurately do these models fit the observed data? The models make both cross-sectional errors (e.g., a three-factor model does not produce a month-t term structure that fits exactly each yield observed at t) and forecast errors. Here I examine the magnitude of these errors.

Three conclusions stand out from the others that are drawn from this evidence. First, differences in forecast accuracy across all models with at least three factors are very small—at most a few basis points. Second, maximum likelihood accommodates the Sharpe ratio constraint by giving up a small amount of in-sample forecast accuracy. There is no loss in cross-sectional accuracy. Third, for a fixed Sharpe ratio constraint, a model with $n + 1$ factors has smaller cross-sectional errors than an n factor model and it has slightly larger forecast errors.

This analysis is restricted to the four unconstrained models and the three-factor, four-factor, and five-factor models estimated subject to the constraint $c = 0.25$. I construct cross-sectional errors by subtracting from observed yields the yields implied by the filtered state vector produced by the Kalman filter. I construct forecast errors at the three-month and twelve-month horizons. The filtered state vector, combined with estimated factor dynamics, produces expected state vectors at these horizons. The corresponding implied bond yields are subtracted from actual yields observed at the future dates. Forecast errors are also constructed using the assumption that yields at all maturities follow random walks.

The relevant information is in Table 7. To summarize the information in forecast errors, I focus on forecast errors in level, slope, and curvature. These are defined following Duffee (2009) as the five-year yield, the five-year yield less the three-month yield, and the two-year yield less the average of the three-month and five-year yields. The two-factor model does a poor job fitting the cross section, with a RMSE exceeding 15 basis points of annualized yields. It also does a relatively poor job forecasting curvature. For example, the RMSE at a three-month horizon is 28 basis points. No other model has a corresponding RMSE greater than 22 basis points.

Aside from the two-factor model, forecast accuracies for all models are within a couple of basis points. For example, at the twelve-month horizon, RMSEs for the level of the term structure range from 79 to 82 basis points. The range at the same horizon for slope (curvature) is 77 to 80 basis points (25 to 27 basis points). Differences in cross-sectional accuracy are also on the order of a few basis points, but they are proportionally much larger. For example, the three-factor models have cross-sectional RMSEs of 7.5 basis points, which is more than 1.5 times the cross-sectional RMSEs of the four-factor models.

For a given number of factors, the unconstrained and constrained models have identical cross-sectional accuracy and slightly different forecast accuracy. The five-factor model has the largest differences. At the three-month horizon the constrained five-factor model has RMSEs for level, slope and curvature that are about one basis point higher than the unconstrained model's RMSEs. At the twelve-month horizon the differences are about three basis points.

Finally, note that among the constrained models, forecasts from the three-factor model are slightly more accurate than those from the four-factor model. These, in turn, are typically slightly more accurate than those from the five-factor model, although this pattern does not hold for three-month-ahead forecasts of the level. All of the differences are on the order of one or two basis points. These economically small differences are statistically large. Inspection of the log-likelihoods reported in Panel B of Table 2 for these models shows that the log-likelihood of the four-factor model exceeds that of the three-factor model by 560. The difference between the five-factor and four-factor models is about 100.

It is worth emphasizing that this analysis is concerned only with in-sample errors. Analysis of out-of-sample errors is considerably more complicated because it requires constrained ML estimation on data samples through t, $t+1$, $\ldots$, T. Whether differences in forecast accuracy are also small out-of-sample is an open question.

7 Concluding comments

This paper explores the role of Sharpe ratios in term structure models. Constraints on Sharpe ratios are imposed without specific parametric restrictions. This approach allows us to determine which features of the models remain after imposing a reasonableness standard on Sharpe ratios.

The next step is to transform the results into parametric restrictions. A natural choice, given the evidence here, is a four-factor model. The first two principal components of yields (level and slope) are priced factors, while the other two factors are unpriced. Risk compensation varies with the slope and fourth factor. The model's mean short rate should be fixed to a relatively high level. Whether such a model generates reasonable Sharpe ratios without explicitly constraining them is the subject of future work.

8 Appendix

8.1 Cochrane and Piazzesi (2005)

Cochrane and Piazzesi (2005) use five forward rates to predict excess log annual returns to Treasury bonds. The forward rates are for maturities between i and $i+1$ years, for

$i = 1, \ldots, 4$. The returns are for zero-coupon bonds with maturities of two through five years, in excess of the return to a one-year zero-coupon bond. Their appendix constructs a no-arbitrage model that fits exactly the coefficients of the predictive regressions. Here I follow their procedure and calculate various Sharpe ratios implied by the model.

The model's factors are log bond prices with maturities from one to five years. Consider the regression in which log bond prices are predicted with year-ago log bond prices. The notation of the stacked regressions is

$$p_{t+12} = \mu + Kp_t + \Sigma\epsilon_{t+12}, \quad \epsilon_{t+12} \sim MVN(0, I). \tag{A1}$$

Cochrane and Piazzesi treat these regressions as a VAR(1) in annual data that is estimated with overlapping observations. The regressions contain the same information as regressions of log excess returns from t to $t + 12$ on five month-t forward rates. In the body of their paper, the authors focus on a single linear combination of forward rates that forecasts. No parameter restrictions are imposed on (A1), thus the estimated model is equivalent to allowing different linear combinations to forecast excess returns to different bonds.

Cochrane and Piazzesi construct equivalent-martingale dynamics of (A1) that are consistent with the use of log prices as factors. These dynamics are not pinned down completely by no-arbitrage, thus they focus on the dynamics that minimize the variance of the stochastic discount factor. Define Q, R, and V as

$$Q = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \qquad R = \begin{pmatrix} -1 & 1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 1 \end{pmatrix}, \qquad V = \Sigma\Sigma'.$$

In the notation of Section 4, the parameters of risk compensation are

$$\lambda_0 = VQ' \left(QVQ'\right)^{-1} \left(Q\mu + \frac{1}{2}\, Q \operatorname{diag}(V)\right),$$

$$\lambda_1 = VQ' \left(QVQ'\right)^{-1} \left(QK - R\right).$$

The parameters λ_0 and λ_1, combined with the panel of factors p_t, allow calculation of compensation for factor risk using equation (18) in the paper's text. Maximum Sharpe ratios can also be constructed.

I estimate the dynamics (A1) using overlapping monthly observations of log bond prices. One sample period is January 1965 through December 2003, which matches Cochrane and

Piazzesi. Another is 1972 through 2008, which is (almost) the sample period studied in this paper. I then calculate Sharpe ratios. Table A1 reports sample means of maximum Sharpe ratios (using both simple and log returns). It also reports sample means of maximum Sharpe ratios for portfolios constructed with excess simple returns to bonds with maturities of two to fifteen years. Finally, the table reports unconditional maximum Sharpe ratios for the same bond portfolios. Figure A1 displays the time series of all three maximum conditional Sharpe ratios for the 1965–2003 period.

The properties of Sharpe ratios in the table are consistent with those of the unconstrained five-factor model reported in the paper's Table 2. In both periods, the sample mean Sharpe ratios using simple returns exceed 10^{16}. For both samples, the maximum unconditional Sharpe ratios using bonds with maturities no greater than 15 years comfortably exceeds the sample unconditional Sharpe ratio for stocks over the same period.[8] (Sharpe ratios for the aggregate stock market are not reported in the Table.)

Cochrane and Piazzesi emphasize the high R^2s of regressions that predict excess log returns to bonds. However, the estimated dynamics of (A1) imply much greater predictability of excess returns to strategies that buy a bond and short two bonds with maturities on either side of the purchased bond. Denote the excess log return to a k-year bond from month t to month $t+1$ by $xr_{t,t+12}^{(k)}$. The regressions are

$$xr_{t,t+12}^{(k)} - \frac{1}{2}\left(xr_{t,t+12}^{(k+1)} + xr_{t,t+12}^{(k-1)}\right) = b_{k,0} + b'_{k,1} f_t + \epsilon_{t,t+12}^{(k)}, \quad k = 3, 4,$$

where f_t is a vector of the five forward rates used by Cochrane and Piazzesi. For the sample 1965 through 2003, the R^2s of these two regressions are 0.52 and 0.61. For the more recent sample, the corresponding R^2s are 0.51 and 0.58. The fitted no-arbitrage model, because it reproduces the OLS estimates of (A1), exhibits variations in prices of risk that are necessary to rationalize these R^2s.

8.2 Sharpe ratio intuition in a one-factor setting

Section 6 documents that when Sharpe ratios are calculated using simple returns, conditional monthly Sharpe ratios of bonds can be substantially less than maximum conditional monthly Sharpe ratios. This appendix uses a one-factor setting to explain this result.

Briefly, both bond prices and the stochastic discount factor are log-normally distributed over finite horizons. A log-normally distributed variable is a nonlinear function of shocks to the state, where the magnitude of the nonlinearity is increasing in volatility. Volatilities of log

[8] Annual simple returns to the stock market are constructed using CRSP value-weighted returns. Excess returns are constructed by subtracting the simple return to a one-year Treasury bond.

bond returns are sufficiently small over monthly horizons that nonlinearities are negligible. Thus when the volatility of the log SDF is very high, bond returns (close to linear in the state) and the SDF (highly nonlinear) are only weakly correlated.

8.2.1 A one-factor model in discrete time

Here the short rate is the only factor. It follows a first-order autoregression

$$r_{t+1} = \mu + K r_t + \sigma \epsilon_{t+1}, \quad \epsilon_{t+1} \sim N(0,1).$$

The log SDF is

$$m_{t+1} = -r_t - \frac{1}{2}\Lambda_t^2 - \Lambda_t \epsilon_{t+1}.$$

It is helpful to write the SDF as the product of the inverse of the gross risk-free return and the Radon-Nikodym derivative,

$$M_{t+1} = R_{f,t+1}^{-1}\xi_{t+1}, \quad \xi_{t+1} = \exp\left(-\frac{1}{2}\Lambda_t^2 - \Lambda_t \epsilon_{t_1}\right) \tag{44}$$

The log return to a k-period bond from t to $t+1$ is normally distributed and given by

$$p_{t+1}^{(k-1)} - p_t^{(k)} = r_t + B_{k-1}\sigma\Lambda_t - \frac{1}{2}B_{k-1}^2\sigma^2 + B_{k-1}\sigma\epsilon_{t+1}.$$

With a single factor, log returns to all bonds are perfectly correlated, and their correlation with the log SDF is either one or minus one. Each bond's absolute conditional Sharpe ratio using log returns equals the maximum conditional Sharpe ratio (using log returns),

$$\theta_t = |s_t^{(k)}| = |\Lambda_t|. \tag{45}$$

The independence between maturity k and the Sharpe ratio does not quite carry over to bonds' absolute conditional Sharpe ratios using simple returns, which are

$$|S_t^{(k)}| = \frac{|1 - \exp\left(-B_{k-1}\sigma\Lambda_t\right)|}{\left(\exp\left(B_{k-1}^2\sigma^2\right) - 1\right)^{1/2}}. \tag{46}$$

When the terms inside the exp functions in (46) are sufficiently close to zero, applying the approximation $(\exp(x) - 1 \approx x)$ to both the numerator and denominator produces $|\Lambda_t|$. Thus for $B_{k-1}\sigma$ close to zero, conditional Sharpe ratios for log bond returns and simple bond returns coincide. By contrast, using simple returns, the maximum conditional Sharpe ratio among fixed-income instruments is

$$\Theta_t = \left(\exp(\Lambda_t^2) - 1\right)^{1/2}. \tag{47}$$

Armed with this machinery, it is easy to demonstrate that the curvature inherent in the Radon-Nikodym derivative function in (44) drives the wedge between the two maximum Sharpe ratios in (45) and (47). Since log bond returns are homoskedastic, heteroskedasticity in the Radon-Nikodym derivative is necessary to generate time-varying expected excess log bond returns. Because the derivative is bounded below by zero, heteroskedasticity corresponds to time-varying curvature in the relation between realizations of the derivative and realizations of interest-rate shocks.

When the volatility of the derivative is high, and curvature of the derivative function is also high. By contrast, the relation between realized simple bond returns and interest-rate shocks is nearly linear, regardless of the volatility of the Radon-Nikodym derivative. Thus times when the maximum conditional Sharpe ratio is very high are also times when simple bond returns have relatively low correlations with the Radon-Nikodym derivative. Thus bond returns do not attain the maximum Sharpe ratio using simple returns.

I use a parameterization of this one-factor model to illustrate the effect.[9] Panel A of Figure A2 reports the term in brackets in

$$P_{t+1}^{(k)} = \exp\left(A_k + B_k E_t(r_{t+1})\right)\left[\exp\left(B_k \sigma \epsilon_{t+1}\right)\right]$$

for $k = 60$ (a five-year bond). It is a scaled version of function mapping the short-rate shock ϵ_{t+1} to the bond price. Naturally, positive shocks to the short rate lower the bond price. Over the range of the horizontal axis, which is plus and minus two standard deviations of the short-rate shock, the function is nearly linear.

The remaining panels in the figure display the ξ_{t+1} functions for three choices of Λ_t. Panel B is constructed using the unconditional mean of Λ_t, while Panels C and D are constructed using Λ_t's that are minus and plus one (unconditional) standard deviation around the mean.

In Panel B, $\Lambda_t = -0.06$. The positive slope in Panel B indicates that investors prefer payoffs that are unexpectedly high when the short rate jumps up. Thus bonds, on average, have positive expected excess returns to compensate investors for the negative covariance between

[9]The parameterization is chosen to reproduce high mean conditional Sharpe ratios, which cannot be done in a one-factor model with realistic parameters. The parameters are $\mu = 0.000015$, $K = 0.997$, $\sigma = 0.0087$, $\lambda_0 = -0.001$, and $\lambda_1 = 0.1$. Interpreting a period as a month, these parameters imply a mean short rate of 6 percent/year, a mean five-year bond yield of 7.65 percent/year, a monthly standard deviation of 87 basis points (three percent/year). In addition, a whopping 63 percent of the variation in monthly excess log bond returns are predictable.

their payoffs and the Radon-Nikodym derivative. Because the Radon-Nikodym derivative function is almost linear in the short-rate shock, the correlation between realizations of the bond price and the derivative is close to perfect (-0.995). Therefore the bond's conditional Sharpe ratio using simple returns is almost identical to the maximum conditional Sharpe ratio using simple returns. The former is 0.0575 and the latter is 0.0578.

In Panel C, $\Lambda_t = -1.35$. The strong curvature of the Radon-Nikodym derivative results in a correlation with the bond price of only -0.56. Therefore, although the maximum Sharpe ratio using simple returns is 2.28, the bond's conditional Sharpe ratio using simple returns is only 1.27. In Panel D, the price of risk changes sign; $\Lambda_t = 1.23$. Bonds are hedges. The correlation between the bond's price and the Radon-Nikodym derivative is 0.69. The bond's absolute conditional Sharpe ratio using simple returns is 1.30, while the maximum conditional Sharpe ratio using simple returns is 1.89.

I use Monte Carlo simulations to calculate the mean of the period-t maximum conditional Sharpe ratios using log and simple returns, as well as the mean of the absolute Sharpe ratio for simple returns to five-year bond. The mean maximum conditional Sharpe ratio using log returns is 1.030. The mean absolute conditional Sharpe ratio for the five-year bond is 1.033. The mean maximum conditional Sharpe ratio using simple returns is 2,692.3. Again, this example is chosen to illustrate what happens when conditional Sharpe ratios can reach unreasonably high values.

8.2.2 A continuous-time model

In the discrete-time model, bonds do not attain the maximum conditional Sharpe ratio using simple returns because they are not perfectly correlated with the stochastic discount factor. The same discrete-time result holds in a continuous-time model. An illustration using the Vasicek (1977) model is sufficient. The short rate follows the process

$$dr = (\mu - kr_t)dt + \sigma dW.$$

The dynamics of the state-price density are

$$\frac{d\pi}{\pi} = -r_t dt - \Lambda dW.$$

All risky financial instruments attain the same absolute instantaneous Sharpe ratio (with the usual abuse of notation), $|\Lambda|\sqrt{dt}$. There is no distinction between Sharpe ratios of log returns and simple returns.

Sharpe ratios for finite-horizon simple returns differ across instruments. Put differently, Sharpe ratios at finite horizons are partly determined by the choice of dynamic trading strategy. depend on dynamic trading strategies, Compute excess returns from t to s, $s > t$, by subtracting the return to a bond that matures at s. Then the Sharpe ratio for the simple return to asset i is

$$\frac{R_{i,t,s} - R_{f,t,s}}{\sqrt{\text{Var}_t\left(R_{i,t,s}\right)}} = -R_{f,t,s}\text{Cor}_t\left(R_{i,t,s}, \frac{\pi_s}{\pi_t}\right)\sqrt{\text{Var}_t\left(\frac{\pi_s}{\pi_t}\right)}.$$

A security with a single payoff at time s of $a - b(\pi_s/\pi_t)$ has the maximum Sharpe ratio for this horizon.

Computation of the variance of π_s/π_t is more difficult in the continuous-time model than in the discrete-time model because the short rate varies over the interval (t, s). When this variation is small relative to the variation in the Radon-Nikodym derivative (equivalently, when σ is small relative to Λ), the maximum simple-return Sharpe ratio in the continuous-time model approaches the counterpart of (47). It is

$$\Theta_t \approx \left(\exp(\Lambda^2(s-t)) - 1\right)^{1/2}.$$

Sharpe ratios for simple bond returns depend on the model's parameters. For plausible choices of σ, these Sharpe ratios are close to $\Lambda\sqrt{s-t}$ for $s - t$ on the order of a month. At this horizon, nonlinearities in bond returns are negligible.

References

Ang, Andrew, and Monika Piazzesi, 2003, A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables, *Journal of Monetary Economics* 50, 745-787.

Campbell, John Y., 2003, Consumption-based asset pricing, in George Constantinides, Milton Harris, and Rene Stulz, Eds., *Handbook of the Economics of Finance, Volume 1B*, North-Holland, Amsterdam.

Campbell, John Y., and Luis M. Viceira, 2001, Who should buy long-term bonds? *American Economic Review* 91, 99-127.

Campbell, John Y., and Luis M. Viceira, 2005, The term structure of the risk-return trade-off, *Financial Analysts Journal* 61, 34-44.

Cochrane, John H., and Monika Piazzesi, 2005, Bond risk premia, *American Economic Review* 95, 139-160.

Cochrane, John H., and Monika Piazzesi, 2008, Decomposing the yield curve, Working paper, Chicago GSB.

Cochrane, John H., and Jesus Saa-Requejo, 2000, Beyond arbitrage: good-deal asset price bounds in incomplete markets, *Journal of Political Economy* 108, 79-119.

Duffee, Gregory R., 1996, Idiosyncratic variation of Treasury bill yields, *Journal of Finance* 51, 527-551.

Duffee, Gregory R., 2002, Term premia and interest rate forecasts in affine models, *Journal of Finance* 57, 405-443.

Duffee, Gregory R., 2008, Information in (and not in) the term structure, working paper, Johns Hopkins.

Duffee, Gregory R., 2009, Forecasting with the term structure: the role of no-arbitrage, working paper, Johns Hopkins.

Duffie, Darrell, and Rui Kan, 1996, A yield-factor model of interest rates, *Mathematical Finance* 6, 379-406.

Fama, Eugene F., and Robert R. Bliss, 1987, The information in long-maturity forward rates, *American Economic Review* 77, 680-692.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3-56.

Gibbons, Michael R., Stephen A. Ross, and Jay Shanken, 1989, A test of the efficiency of a given portfolio, *Econometrica* 57, 1121-1152.

Gurkaynak, Refet S., Brian Sack, and Jonathan H. Wright, 2006, The U.S. Treasury yield curve: 1961 to the present, working paper, Federal Reserve Board.

Hansen, Lars Peter, and Ravi Jagannathan, 1991, Implications of security market data for models of dynamic economies, *Journal of Political Economy* 99, 225-262.

Joslin, Scott, Marcel Priebsch, and Kenneth J. Singleton, 2009, Risk premium accounting in macro-dynamic term structure models, Working paper, Stanford GSB.

Luttmer, Erzo G. J., 1996, Asset pricing in markets with frictions, *Econometrica* 64, 1439-1467.

MacKinlay, A. Craig, 1995, Multifactor models do not explain deviations from the CAPM, *Journal of Financial Economics* 38, 3-28.

Ross, Stephen A., 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13, 341-360.

Sangvinatsos, Antonios, and Jessica A. Wachter, 2005, Does the failure of the expectations hypothesis matter for long-term investors?, *Journal of Finance* 60, 179-230.

Sharpe, William F., 1966, Mutual fund performance, *Journal of Business* 39, 119-138.

Vasicek, Oldrich, 1977, An equilibrium characterization of the term structure, *Journal of Financial Economics* 5, 177-188.

Table 1. Sample unconditional Sharpe ratios, 1952 through 2008

The table reports sample means, standard deviations, and unconditional Sharpe ratios of excess nominal returns to Treasury bond portfolios and the aggregate stock market. The sample is January 1952 through December 2008. Monthly excess returns are constructed by subtracting the return to one-month Treasury bills as calculated by Ibbotson Associates. At the quarterly horizon, the return to three-month Treasury bills is used. The table also reports maximum Sharpe ratios for unconstrained positions in the portfolios of Treasury bonds. Means and standard deviations are in percent per horizon (monthly or quarterly).

Portfolio (m in months)	Monthly horizon Mean	Monthly horizon Std dev	Monthly horizon Sharpe	Quarterly horizon Mean	Quarterly horizon Std dev	Quarterly horizon Sharpe
$0 < m \leq 6$	0.039	0.139	0.28	0.061	0.263	0.23
$6 < m \leq 12$	0.062	0.344	0.18	0.130	0.673	0.19
$12 < m \leq 24$	0.088	0.613	0.14	0.206	1.190	0.17
$24 < m \leq 36$	0.112	0.936	0.12	0.281	1.777	0.16
$36 < m \leq 48$	0.126	1.171	0.11	0.323	2.226	0.14
$48 < m \leq 60$	0.123	1.376	0.09	0.306	2.594	0.12
$60 < m \leq 120$	0.143	1.672	0.09	0.361	3.113	0.12
Stock market	0.493	4.292	0.11	1.483	7.844	0.19
Max using final six bond portfolios	–	–	0.23	–	–	0.26

Table 2. Maximum Sharpe ratios of estimated term structure models

Gaussian dynamic term structure models are estimated with maximum likelihood using the Kalman filter. The data are a panel of 446 months of Treasury yields from November 1971 through December 2008. Conditional maximum monthly Sharpe ratios for both log returns and simple returns are calculated for each month, assuming complete fixed-income markets up to a ten-year maturity. Ratios are also calculated for simple returns assuming a complete bond market. The table reports sample means of these monthly Sharpe ratios (top number) as well as model-implied population means (bottom number). For a complete bond market, the table also reports the population unconditional maximum monthly Sharpe ratio using simple returns.

Panel A. Unconstrained models

		Bond market, unconditional	Mean of conditional maximum ratios		
Number		maximum ratio,	Fixed-income market		Bond market
of factors	log like	simple returns	log returns	simple returns	simple returns
2	28502.52	0.116	0.200	0.203	0.201
			0.181	0.184	0.182
3	29987.21	0.192	0.331	0.352	0.337
			0.310	0.326	0.316
4	30562.39	0.337	0.556	0.784	0.578
			0.590	0.727	0.613
5	30675.15	0.561	1.762	4.9×10^{30}	2.804
			2.339	4.2×10^{31}	3.712

Panel B. Constraint on sample mean of max Sharpe ratios

		Bond market, unconditional	Mean of conditional maximum ratios		
Number		maximum ratio,	Fixed-income market		Bond market
of factors	log like	simple returns	log returns	simple returns	simple returns
3	29985.18	0.180	0.250	0.259	0.253
			0.265	0.274	0.268
4	30544.26	0.180	0.250	0.264	0.252
			0.310	0.326	0.312
5	30641.82	0.166	0.250	0.269	0.251
			0.336	0.357	0.338

Table 3. Properties of yields implied by estimated term structure models

Gaussian dynamic term structure models are estimated with maximum likelihood using the Kalman filter. The data are a panel of 446 months of Treasury yields from November 1971 through December 2008. Constrained models have a restriction on the model's conditional maximum Sharpe ratios. For each estimated model, the table reports implied unconditional means, one-month-ahead standard deviations, and unconditional standard deviations of yields. Unconditional standard deviations are the top numbers and one-month-ahead standard deviations are the bottom numbers. Sample means and standard deviations are reported in the final row. Yields are expressed in percent per year.

		Means				Standard deviations			
Number		Maturity (months)				Maturity (months)			
of factors	Constrained	1	3	60	120	1	3	60	120
2	N	4.95	5.02	6.08	6.53	3.27	3.24	2.88	2.58
						0.58	0.56	0.37	0.33
3	N	4.86	4.96	6.02	6.51	3.19	3.19	2.87	2.62
						0.59	0.56	0.39	0.32
4	N	4.63	4.81	5.89	6.39	3.29	3.30	2.99	2.72
						0.61	0.56	0.38	0.32
5	N	4.28	4.91	5.99	6.48	3.79	3.25	2.93	2.67
						1.29	0.56	0.38	0.32
3	Y	6.31	6.41	7.44	7.85	3.84	3.84	3.44	3.15
						0.60	0.57	0.39	0.32
4	Y	8.09	8.19	9.13	9.44	4.16	4.13	3.46	3.14
						0.62	0.58	0.39	0.33
5	Y	9.19	9.28	10.09	10.31	4.42	4.40	3.80	3.44
						0.64	0.58	0.39	0.33
Data sample		-	5.93	7.03	7.46	-	3.03	2.66	2.42

Table 4. Properties of the term structure implied by estimated term structure models

Gaussian dynamic term structure models are estimated with maximum likelihood using the Kalman filter. The data are a panel of 446 months of Treasury yields from November 1971 through December 2008. Constrained models have a restriction on the model's conditional maximum Sharpe ratios. For each estimated model, the table reports the population serial correlation of the ten-year yield at the ten-year horizon. It also reports unconditional means and standard deviations of conditional monthly Sharpe ratios (log returns) for bonds of various maturities. Yields are expressed in percent per year. Standard deviations are in parentheses.

			Conditional Sharpe ratio Maturity (months)		
Number of factors	Constrained	Ten-year correlation	3	60	120
2	N	0.27	0.11	0.09	0.07
			(0.13)	(0.15)	(0.15)
3	N	0.30	0.17	0.09	0.08
			(0.19)	(0.20)	(0.22)
4	N	0.30	0.29	0.09	0.08
			(0.38)	(0.24)	(0.27)
5	N	0.28	0.75	0.11	0.09
			(1.76)	(0.31)	(0.31)
3	Y	0.46	0.17	0.08	0.07
			(0.15)	(0.16)	(0.18)
4	Y	0.36	0.17	0.07	0.06
			(0.21)	(0.18)	(0.20)
5	Y	0.50	0.15	0.06	0.05
			(0.22)	(0.20)	(0.21)

Table 5. Excess return predictability implied by estimated term structure models

Gaussian dynamic term structure models are estimated with maximum likelihood using the Kalman filter. The data are a panel of 446 months of Treasury yields from November 1971 through December 2008. Constrained models have a restriction on the model's conditional maximum Sharpe ratios. For each estimated model, the table summarizes population values of the predictability of excess log returns to a ten-year zero-coupon bond. Monthly and annual returns are in excess of returns to one-month and one-year Treasury bonds respectively. The "monthly factor" is the first principal component of conditional expectations of monthly excess log returns to bonds with maturities of $2, 3, \ldots, 10$ years.

Number of factors	Constrained	Excess return standard dev Month	Excess return standard dev Ann	R^2 Month	R^2 Ann	Fraction explained by 'monthly factor' Month	Fraction explained by 'monthly factor' Ann
2	N	0.033	0.109	0.022	0.143	0.997	0.998
3	N	0.033	0.105	0.046	0.190	0.998	0.845
4	N	0.033	0.108	0.069	0.200	0.996	0.696
5	N	0.034	0.109	0.089	0.190	0.992	0.483
3	Y	0.033	0.104	0.031	0.147	0.998	0.841
4	Y	0.033	0.110	0.037	0.195	0.999	0.778
5	Y	0.033	0.109	0.042	0.141	0.999	0.641

Table 6. Principal components decomposition of a four-factor model

A four-factor gaussian dynamic term structure models is estimated with maximum likelihood using the Kalman filter. The data are a panel of 446 months of Treasury yields from November 1971 through December 2008. In estimation, the sample mean of conditional maximum Sharpe ratios is constrained to be no greater than 0.25. The factors are rotated into principal components of yields.

	Factor			
	1	2	3	4
Fraction explained of total variance in yields	0.978	0.021	0.001	0.000
Mean risk compensation per standard dev of excess return				
1-year bond	0.076	0.046	0.003	−0.007
5-year bond	0.093	−0.014	−0.008	0.004
10-year bond	0.098	−0.053	0.020	−0.006
Fraction of predictable variation in monthly log excess returns				
1-year bond	0.061	0.381	0.002	0.556
5-year bond	0.006	0.580	0.015	0.399
10-year bond	0.000	0.548	0.018	0.434
Fraction of predictable variation in annual log excess returns				
2-year bond	0.060	0.823	0.046	0.071
5-year bond	0.011	0.887	0.030	0.072
10-year bond	0.000	0.874	0.028	0.097

Table 7. The accuracy of estimated term structure models

Gaussian dynamic term structure models are estimated with maximum likelihood using the Kalman filter. The data are a panel of 446 months of Treasury yields from November 1971 through December 2008. Constrained models have a restriction on the model's conditional maximum Sharpe ratios. For each estimated model, the table reports root mean squared cross-sectional and forecast errors. The cross-sectional error is the sample mean, across all observed bonds and dates, of the squared difference between the actual yield and the yield implied by the Kalman filter. The forecast errors are for the five-year yield (level), the five-year yield less the three-month yield (slope), and the two-year yield less the average of the three-month and five-year yields (curvature). The table also reports the root mean squared error for the assumption that all yields follow random walks. Errors are expressed in annualized percentage points.

Number of			Three-month horizon			Twelve-month horizon		
factors	Constrained	Cross section	Level	Slope	Curve	Level	Slope	Curve
2	N	15.3	60	61	28	81	79	29
3	N	7.5	58	61	22	79	77	26
4	N	4.8	58	61	22	79	77	26
5	N	4.1	57	60	21	79	77	25
3	Y	7.5	59	61	22	79	78	26
4	Y	4.8	59	62	22	81	80	27
5	Y	4.1	58	62	22	82	80	27
Random walk		-	60	66	23	82	89	30

Table A1. Maximum Sharpe ratios implied by the model of Cochrane and Piazzesi

Bond yields are observed at a monthly frequency. Overlapping observations are used to construct an annual frequency model. Conditional maximum annual Sharpe ratios for both log returns and simple returns are calculated for each month, assuming complete fixed-income markets. Ratios are also calculated for simple returns assuming a complete bond market up to a fifteen-year maturity. The table reports sample means of these Sharpe ratios (top number) as well as model-implied population means (bottom number). For a complete bond market, the table also reports the population unconditional maximum annual Sharpe ratio using simple returns.

	Bond market, unconditional	Mean of conditional maximum ratios		
Number of	maximum ratio,	Fixed-income market		Bond market
factors	simple returns	log returns	simple returns	simple returns
1965–2003	0.504	1.755	3.4×10^{16}	2.653
		1.954	2.1×10^{8}	2.743
1972–2008	0.447	1.656	1.6×10^{18}	2.175
		1.901	1.9×10^{9}	2.346

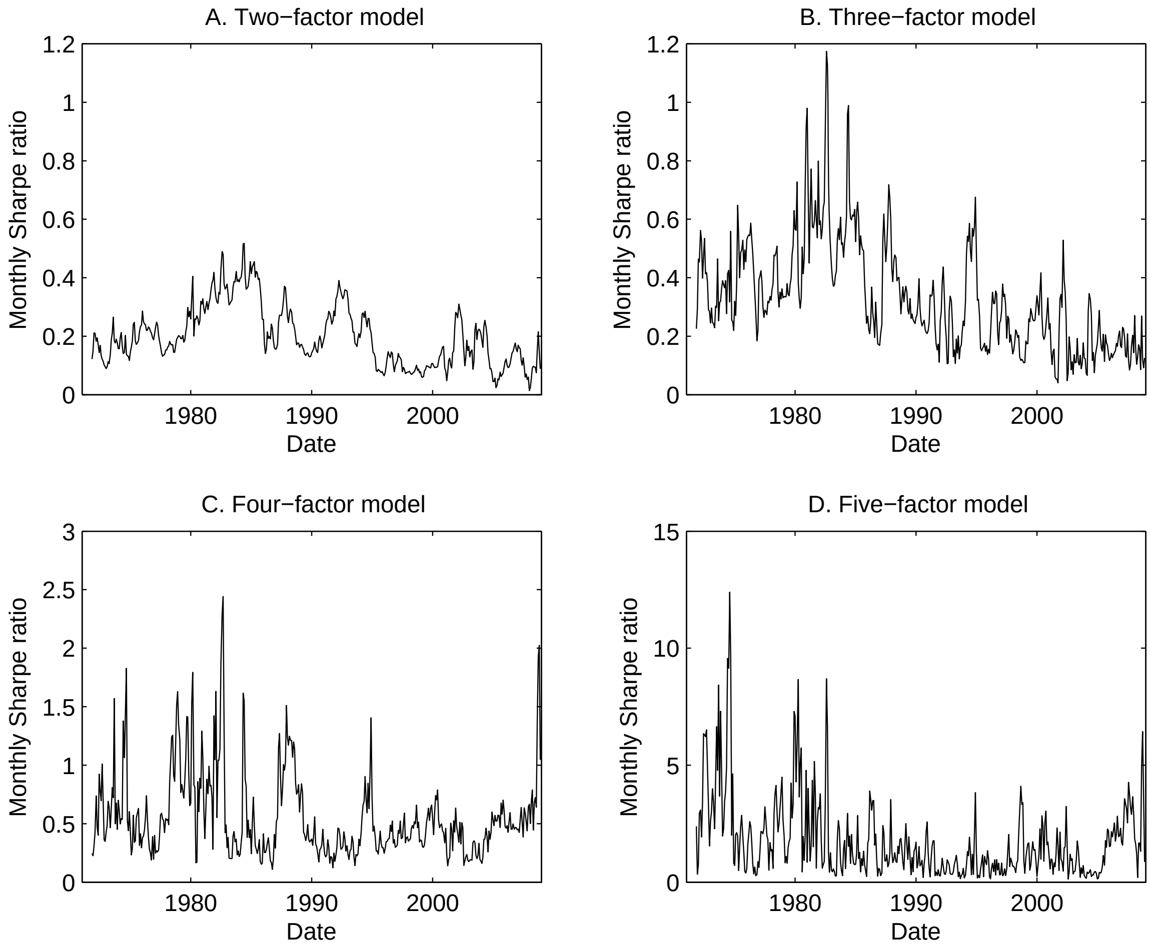

Fig. 1. Conditional maximum Sharpe ratios (log returns) implied by estimates of Gaussian term structure models. The models differ in the number of factors, ranging from two to five. They are estimated on monthly data from November 1971 through December 2008. For each month, the conditional maximum monthly Sharpe ratios, using log returns, are displayed in this figure. Note that the scale of the vertical axes differs across the panels.

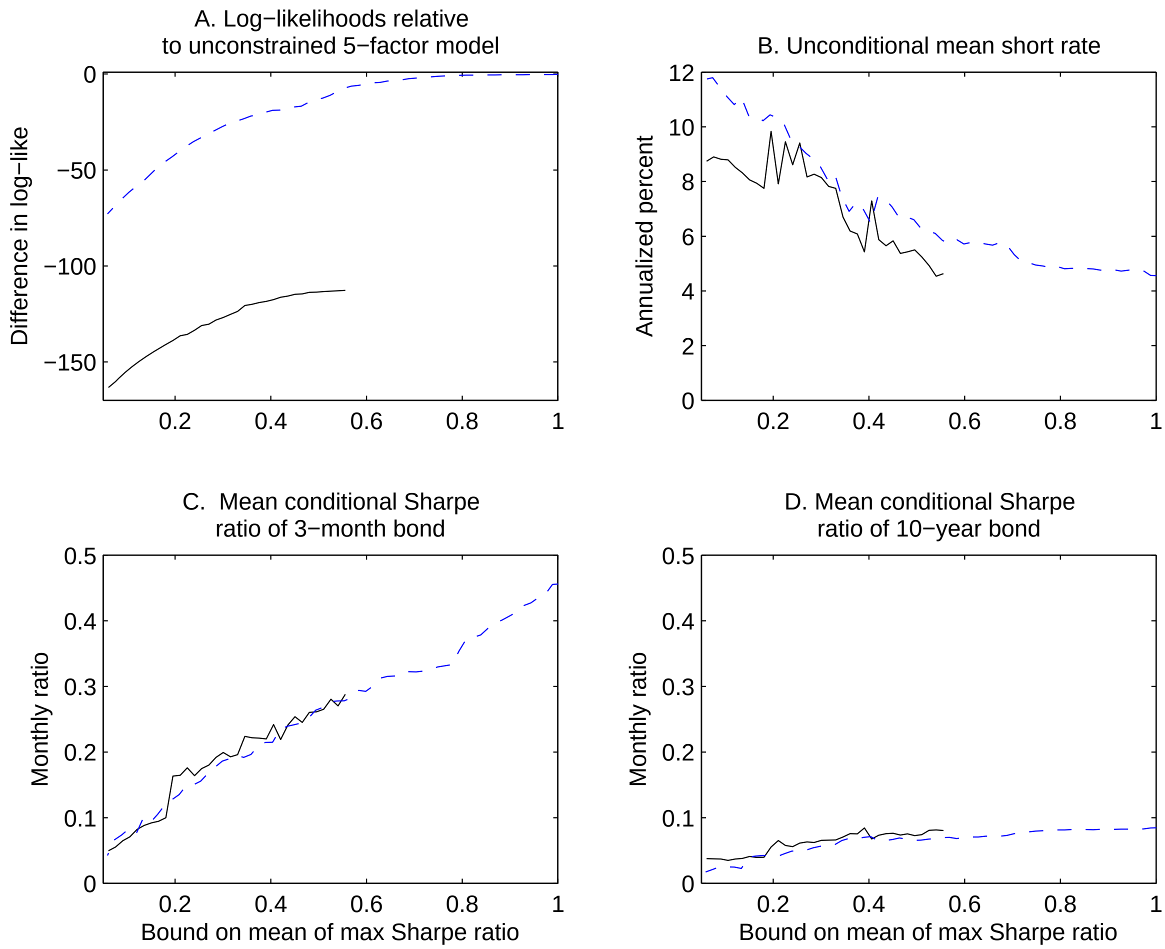

Fig. 2. Characteristics of estimated term structure models. The figure summarizes features of about 100 sets of estimates of Gaussian term structure models. Four-factor and five-factor models are estimated with maximum likelihood, using Treasury yields from November 1971 through December 2008. For a given estimation, model-implied sample means of conditional maximum Sharpe ratios are constrained to not exceed the values on the x-axis. Panel A reports the maximized value of the log-likelihood less the log-likelihood of the unconstrained five-factor model. Panel B reports the unconditional mean yield of a one-month bond. Panels C and D report unconditional means of conditional Sharpe ratios for three-month and ten-year bonds. Solid black lines and dashed blue lines correspond to four-factor and five-factor estimates respectively.

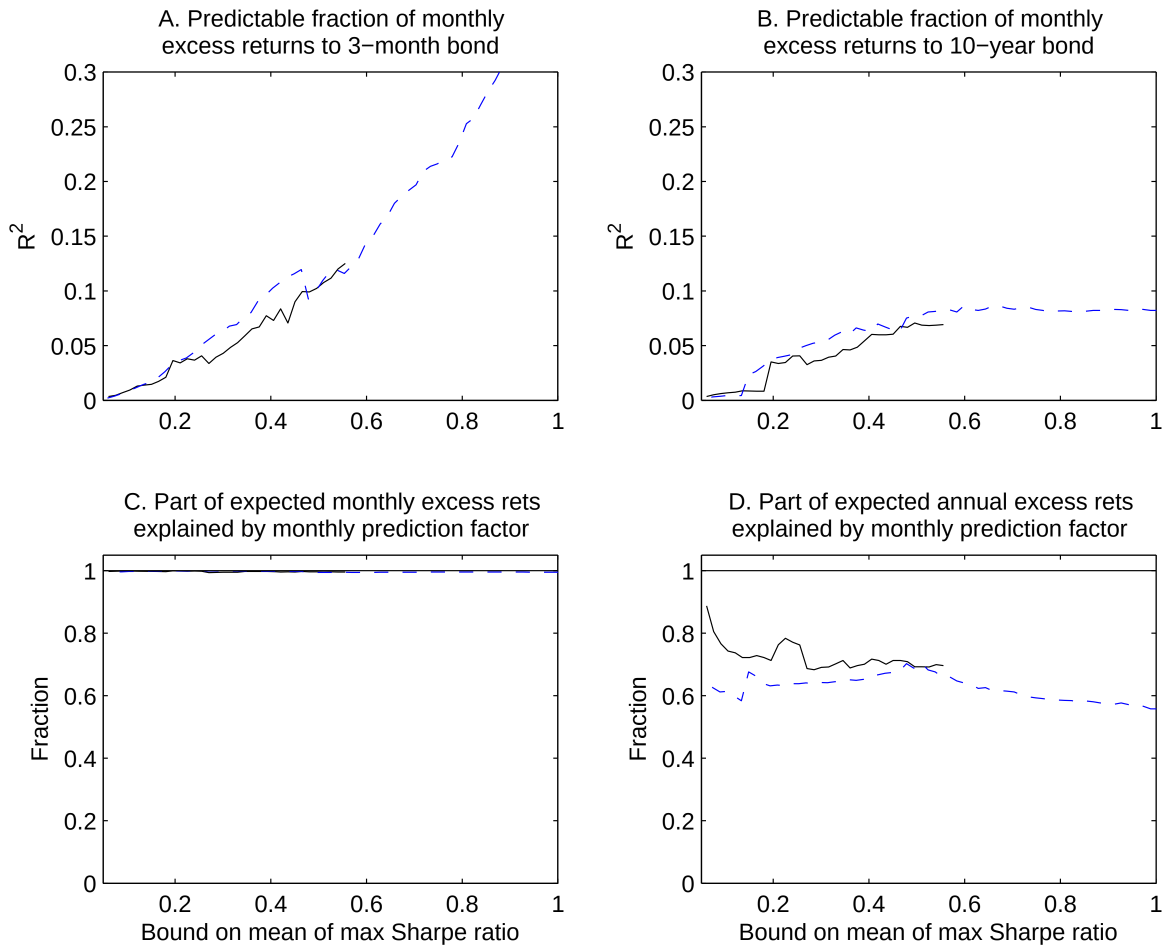

Fig. 3. Additional characteristics of estimated term structure models. Its construction follows that of Fig. 2. Panel A reports the fraction of month $t+1$'s excess log return to a three-month bond that is explained by the month-t state. Panel B reports the same fraction for a ten-year bond. For Panels C and D, the factor rotation corresponds to principal components of the covariance matrix of predictable log excess monthly returns to bonds with maturities ranging from two to ten years. Panel C displays the fraction of this covariance matrix explained by the first factor, denoted the "monthly prediction" factor. Panel D displays the fraction of the predictable log excess annual return to a ten-year bond that is explained by the monthly prediction factor. Solid black lines and dashed blue lines correspond to four-factor and five-factor estimates respectively.

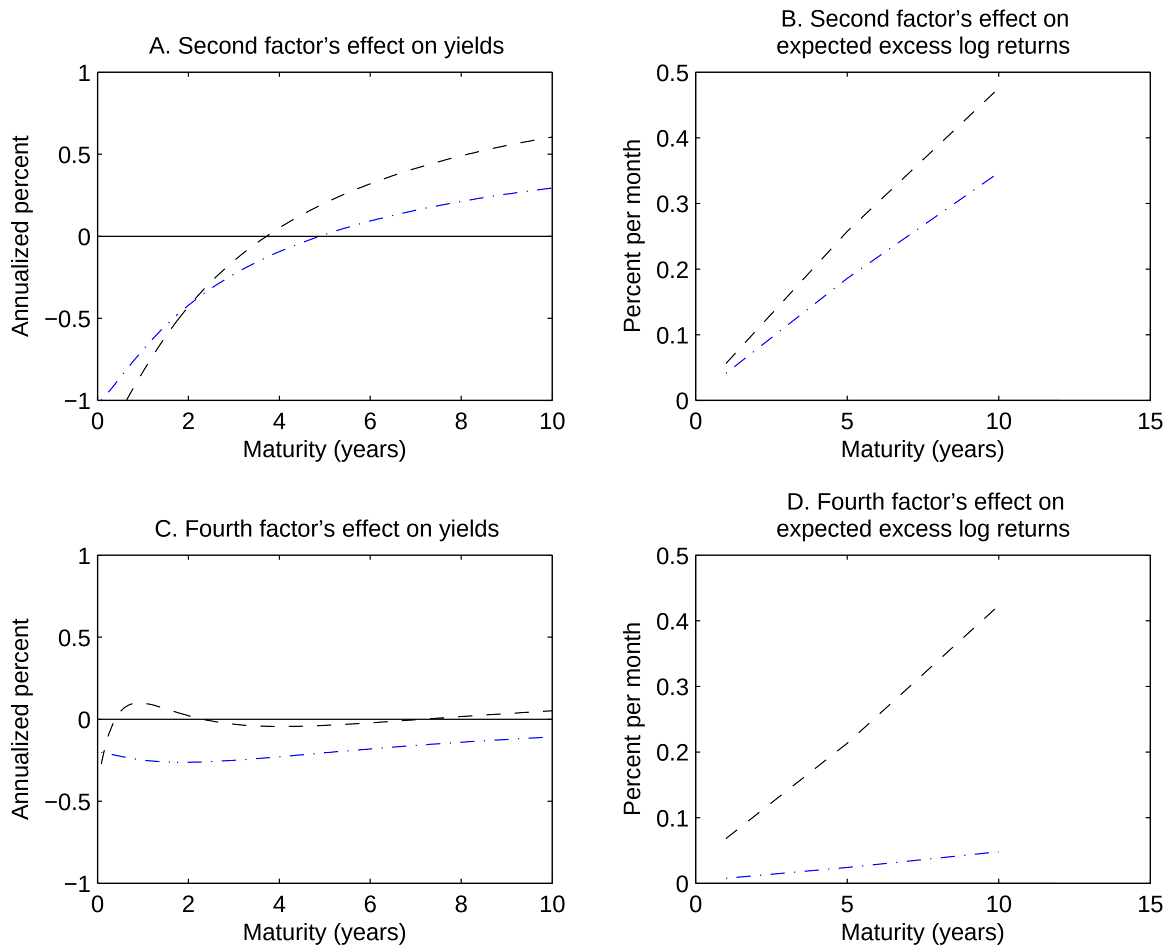

Fig. 4. Some dynamics implied by estimates of a four-factor Gaussian term structure model. In estimation, the sample mean of conditional maximum Sharpe ratios is constrained to be no greater than 0.25. The factors are rotated into principal components of yields. Panels A and C report the reaction of the yield curve to one-standard-deviation changes in the month-t values of the second and fourth factors, respectively. The black dashed line is the month-t effect and the dotted-dashed blue line is month-t expectation of the effect in month $t+12$. Panels B and C report similar information for the conditional expectation of one-month log excess returns to bonds. The dashed black line is the effect on month t's expectation of month $t+1$'s excess returns and the dotted-dashed blue line is the effect on month t's expectation of month $t+13$'s excess returns.

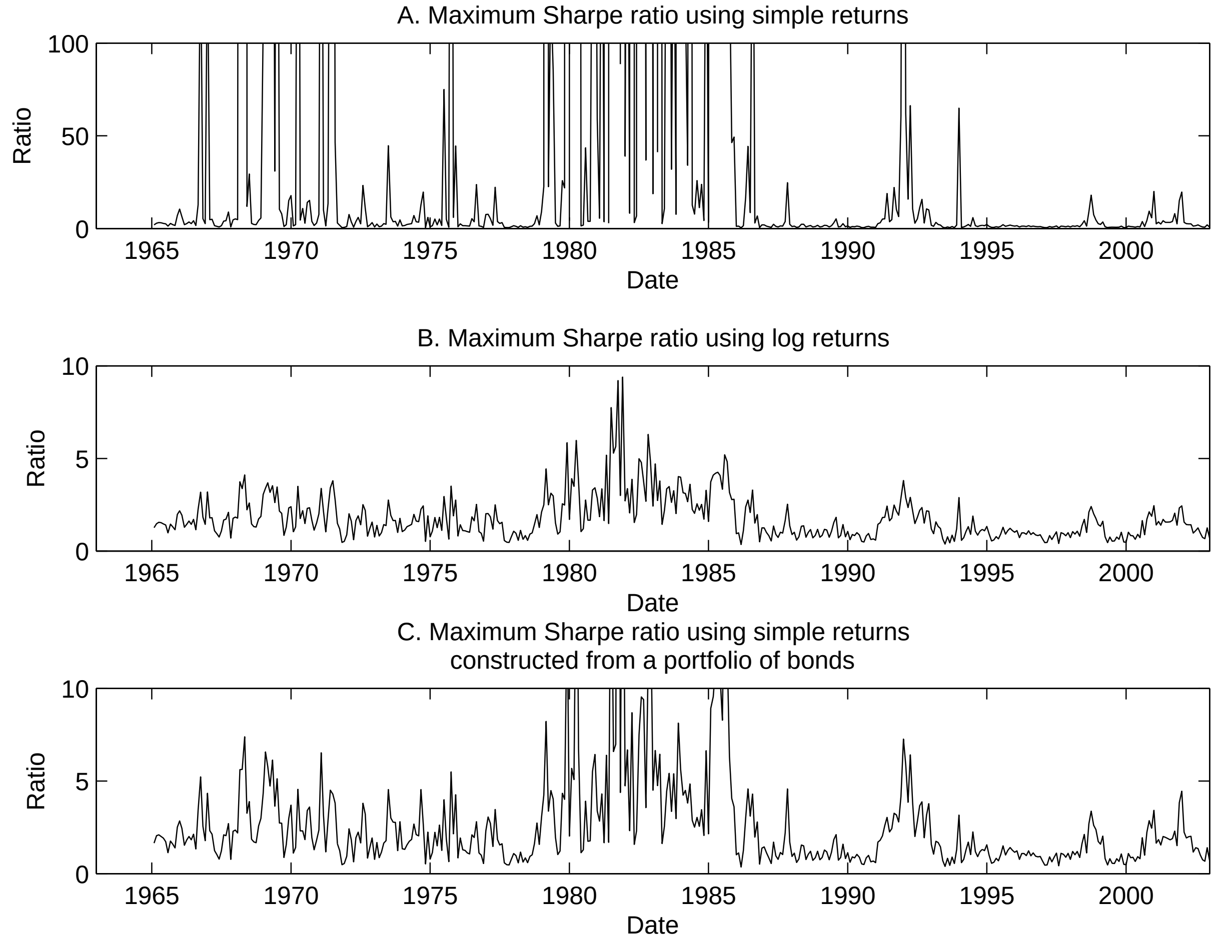

Fig. A1. Conditional maximum annual Sharpe ratios implied by the no-arbitrage model of Cochrane and Piazzesi (2005). The sample period is 1965 through 2003. Panels A and B assume a complete fixed-income market. Panel A uses simple excess returns and Panel B uses log returns. Panel C uses simple excess returns to bonds with maturities from two to fifteen years.

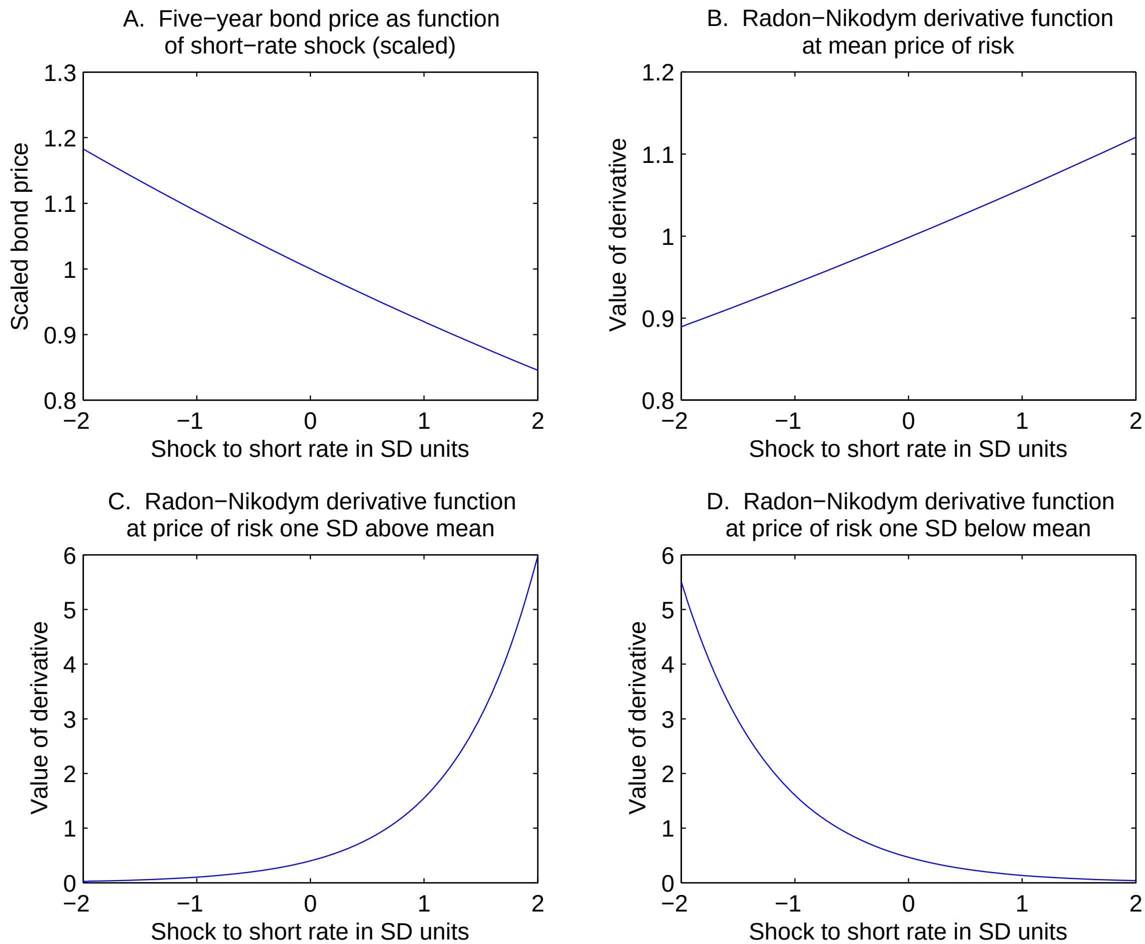

Fig. A2. The price of a five-year bond and the Radon-Nikodym derivative expressed as functions of the short-rate shock in a one-factor Gaussian discrete-time term structure model.