THE JOURNAL OF FINANCE • VOL. LX, NO. 1 • FEBRUARY 2005

Does the Failure of the Expectations Hypothesis Matter for Long-Term Investors?

ANTONIOS SANGVINATSOS and JESSICA A. WACHTER*

ABSTRACT

We solve the portfolio problem of a long-run investor when the term structure is Gaussian and when the investor has access to nominal bonds and stock. We apply our method to a three-factor model that captures the failure of the expectations hypothesis. We extend this model to account for time-varying expected inflation, and estimate the model with both inflation and term structure data. The estimates imply that the bond portfolio of a long-run investor looks very different from the portfolio of a mean-variance optimizer. In particular, time-varying term premia generate large hedging demands for long-term bonds.

THE EXPECTATIONS HYPOTHESIS OF INTEREST RATES states that the premium on long-term bonds over short-term bonds is constant over time. According to this hypothesis, there are no particularly good times to invest in long-term bonds relative to short-term bonds, nor are there particularly bad times. Long-term bonds will always offer the same expected excess return.[1]

While the expectations hypothesis is theoretically appealing, it has consistently failed in U.S. postwar data. Fama and Bliss (1987) and Campbell and Shiller (1991), among others, show that expected excess returns on long-term bonds (term premia) do vary over time, and moreover, it is possible to predict excess returns on bonds using observables such as the forward rate or the term spread. This paper explores the consequences of the failure of the expectations hypothesis for long-term investors.

*Sangvinatsos is at the Stern School of Business at New York University. Wachter is at the Wharton School at the University of Pennsylvania and the NBER. The authors are grateful for helpful comments from Utpal Bhattacharya, Michael Brennan, John Campbell, Jennifer Carpenter, Qiang Dai, Ned Elton, Rick Green, Blake LeBaron, Anthony Lynch, Lasse Pedersen, Matthew Richardson, Luis Viceira, an anonymous referee, and seminar participants at Brandeis University, the Norwegian School of Economics and Business, the Norwegian School of Management, University of California at Los Angeles, Indiana University at Bloomington, the Spring 2003 NBER Asset Pricing meeting, and the 2003 CIREQ-CIRANO-MITACS Conference on Portfolio Choice.

[1] The expectations hypothesis, as we refer to it, should be distinguished from the *pure expectations hypothesis*, which states that term premia are not just constant but equal to zero. Cox, Ingersoll, and Ross (1981) examine variants of the pure expectations hypothesis in the context of continuous-time equilibrium theory, and find that they are inconsistent with each other, and that several imply arbitrage opportunities (see, however, Longstaff (2000)). Campbell (1986) shows that these inconsistencies do not occur with the more general expectations hypothesis, which does not require term premia to be zero. In fact, it is the expectations hypothesis, as opposed to the pure expectations hypothesis, which is typically examined in the empirical literature (see Bekaert and Hodrick (2001) for a discussion of recent empirical work testing the expectations hypothesis).

We estimate a three-factor affine term structure model similar to that proposed in Dai and Singleton (2002) and Duffee (2002) that accounts for the fact that excess bond returns are predictable. We then solve for the optimal portfolio for an investor taking this term structure as given. Bond market predictability will clearly affect the characteristics of the mean-variance efficient portfolio, but the consequences for long-horizon investors go beyond this. Merton (1971) shows that when investment opportunities are time-varying, a mean-variance efficient portfolio is generally suboptimal. Long-horizon investors wish to hedge changes in the investment opportunity set; depending on the level of risk aversion, the investor may want more or less wealth when investment opportunities deteriorate than when they improve. As we will show, investors gain by hedging time-variation in the term premia. Thus, the investor's bond portfolio looks different from that dictated by mean-variance efficiency.

Despite the obvious importance of bonds to investors, as well as the strength of the empirical findings mentioned above, recent literature on portfolio choice has focused almost exclusively on predictability in stock returns. As shown by Fama and French (1989) and Campbell and Shiller (1988), the price-dividend ratio predicts excess stock returns with a negative sign. Based on this finding, a number of studies (e.g., Balduzzi and Lynch (1999), Barberis (2000), Brandt (1999), Brennan, Schwartz, and Lagnado (1997), Campbell and Viceira (1999), Liu (1999), and Wachter (2002)) document gains from timing the stock market based on the price-dividend ratio, as well as from hedging time-variation in expected stock returns. One result of this literature is that when investors have relative risk aversion greater than one, hedging demands dictate that their allocation to stock should increase with the horizon. A natural question to ask is whether the same mechanism is at work for bond returns. Just as stock prices are negatively correlated with increases in future risk premia on stocks, bond prices are negatively correlated with increases in future risk premia on bonds.[2] This intuition suggests that time-variation in risk premia would cause the optimal portfolio allocation to long-term bonds to increase with horizon.

In the case where the investor allocates wealth between a long and a short-term bond, we show that this intuition holds. Hedging demands induced by time-variation in risk premia more than double the investor's allocation to the long-term bond. Moreover, we find large horizon effects. The investor with a horizon of 20 years holds a much greater percentage of his wealth in long-term bonds than an investor with a horizon of 10 years. In the case of multiple long-term bonds, the mean-variance efficient portfolio often consists of a long and short position in long-term bonds. This occurs because of the high positive correlation between bonds of different maturities implied by the model and found in the data. Hedging demand induced by time-varying risk premia generally makes the allocation to long-term bonds more extreme. We find that following a myopic strategy and, in particular, failing to hedge time variation in risk premia carries a high utility cost for the investor.

[2] We consider U.S. government bonds that are not subject to default risk. Nonetheless, we use risk premia and term premia interchangeably, as we do not take a stand on the source of the premia.

Our framework generalizes previous studies of portfolio choice when real interest rates vary over time and there is inflation. Brennan and Xia (2002) and Campbell and Viceira (2001) estimate a two-factor Vasicek (1977) term structure model and determine optimal bond portfolios. Both of these studies assume that risk premia on bonds and stocks are constant.[3] Our study also relates to that of Campbell, Chan, and Viceira (2003), who estimate a vector-autoregression (VAR) including the returns on a long-term bond, a stock index, the dividend yield, and the yield spread. Campbell et al. derive an approximate solution to the optimal portfolio choice problem when asset returns are described by the VAR. The advantage of the VAR approach is that it captures predictability in bond and stock returns in a relatively simple way. The disadvantage is that the term structure is not well defined; it is necessary to assume that the investor only has access to those bonds included in the VAR. Moreover, estimating bond returns using a VAR gives up the extra information resulting from the no-arbitrage restriction on bonds, namely that bonds have to pay their (nominal) face value when they mature.

Rather than modeling bond return predictability using a VAR, we follow the affine bond pricing literature (e.g., Dai and Singleton (2000, 2002a), Duffe (2002)) and specify a nominal pricing kernel.[4] The drift and diffusion of the pricing kernel is driven by three underlying factors that follow a multivariate Ornstein-Uhlenbeck process. The price of risk is a linear function of the state variables. Thus, the model is in the "essentially affine" class proposed by Duffee (2002) and shown by Dai and Singleton (2002) to capture the pattern of bond predictability in the data.

As a necessary step to show the implications of affine term structure models for investors, we show how parameters of the inflation process can be jointly estimated with term structure parameters. This joint estimation produces a series for expected inflation that explains 37% of the variance of realized inflation. This result has implications not only for portfolio choice problems, but also for the estimation of term structure models more generally.

The remainder of the paper is organized as follows. Section I describes the general form of an economy where nominal bond prices are affine, and there exists equity and unhedgeable inflation. Section II derives a closed-form solution for optimal portfolio choice when the investor has utility over terminal wealth and over intermediate consumption. When inflation is introduced, the pricing kernel that determines asset prices is not unique; from the point of view of the investor, it is not well defined. As He and Pearson (1991) show, there is a

[3] Other work on bond returns and portfolio choice includes Brennan and Xia (2000) and Sorensen (1999), who assume that interest rates are as in Vasicek (1977), and Liu (1999) and Schroder and Skiadas (1999), who assume general affine dynamics. These studies assume that bonds are indexed, or equivalently, that there is no inflation. Xia (2002) examines the welfare consequences of limited access to nominal bonds under a Vasicek model. Wachter (2003) shows under general conditions that as risk aversion approaches infinity, the investor's allocation approaches 100% in a long-term indexed bond. None of these papers explore the consequences of bond return predictability.

[4] For recent surveys of this literature, see Dai and Singleton (2003) and Piazzesi (2002).

unique pricing kernel that gives the marginal utility process for the investor.[5] We derive a closed-form expression for this pricing kernel when incompleteness results from inflation. This expression holds regardless of the form of the term structure. Section III uses maximum likelihood to estimate the parameters of the process and demonstrates that the model provides a good fit to term structure data and to inflation. Section IV discusses the properties of the optimal portfolio for the parameters we have estimated and calculates utility costs resulting from suboptimal strategies. Section V concludes.

I. The Economy

As in the affine term structure literature, we specify an exogenous nominal pricing kernel. Because our purpose is modeling predictability in excess bond returns and, as Dai and Singleton (2002) and Duffee (2002) show, a Gaussian model is well suited for this purpose, we will assume that all variables are homoscedastic.[6]

Let dz denote a $d \times 1$ vector of independent Brownian motions. Let $r(t)$ denote the instantaneous nominal riskfree rate. We assume that

$$r(X(t), t) = \delta_0 + \delta X(t), \tag{1}$$

where $X(t)$ is an $m \times 1$ vector of state variables that follow the process

$$dX(t) = K(\theta - X(t))\,dt + \sigma_X\,dz(t), \tag{2}$$

under the physical measure. The matrix of loadings on the Brownian motions, σ_X, is $m \times d$, K is $m \times m$, and θ is $m \times 1$. Suppose there exists a price of risk $\bar{\Lambda}(t)$ that is linear in $X(t)$:

$$\bar{\Lambda}(t) = \bar{\lambda}_1 + \bar{\lambda}_2 X(t), \tag{3}$$

where $\bar{\lambda}_1$ is $d \times 1$ and $\bar{\lambda}_2$ is $d \times m$. When $\bar{\lambda}_2 = 0_{d\times m}$, the price of risk is constant and the model is a multi-factor version of Vasicek (1977). Given a process for the interest rate r and the price of risk $\bar{\Lambda}$, the pricing kernel is given by

$$\frac{d\bar{\phi}(t)}{\bar{\phi}(t)} = -r(t)\,dt - \bar{\Lambda}(t)^\top dz. \tag{4}$$

The pricing kernel determines the price of an asset based on its nominal payoff.

[5] Liu and Pan (2003) also associate the pricing kernel in the economy with the pricing kernel for the investor. In Liu and Pan's model markets are complete, so a unique pricing kernel exists.

[6] Fisher (1998) shows that a two-factor Gaussian model can partially replicate the failure of the expectations hypothesis, but does not make comparisons across models. Bansal and Zhou (2002) show that a regime-switching is also successful at capturing the failure of the expectations hypothesis in the data. Ahn, Dittmar, and Gallant (2002) discuss an affine-quadratic class of models which, as Brandt and Chapman (2002) show, is also capable of accounting for the failure of the expectations hypothesis. Extensions of the results in this paper to quadratic models and models with regime shifts will be considered in future work.

In this economy, bond yields are affine in the state variables $X(t)$. Let $P(X(t), t, s)$ denote the price of such a bond maturing at $s > t$. Then P equals the present discounted value of the bond payoff, namely \$1:

$$P(X(t), t, s) = \bar{\phi}(t)^{-1} E_t[\bar{\phi}(s)].$$

As shown by Duffie and Kan (1996), nominal bond prices take the form

$$P(X(t), t, s) = \exp\{A_2(s-t)X(t) + A_1(s-t)\}, \tag{5}$$

where $A_2(\tau)$ and $A_1(\tau)$ solve a system of ordinary differential equations given in Appendix A. Bond yields are given by

$$\begin{aligned} y(X(t), t, s) &= -\frac{1}{s-t} \log P(X(t), t, s) \\ &= -\frac{1}{s-t}(A_2(s-t)X(t) + A_1(s-t)). \end{aligned} \tag{6}$$

The dynamics of bond prices follow from Ito's lemma:

$$\begin{aligned} \frac{dP(t)}{P(t)} = &\left\{ -A_2'(\tau)X(t) - A_1'(\tau) + A_2(\tau)K(\theta - X(t)) + \frac{1}{2} A_2(\tau)\sigma_X \sigma_X{}^\top A_2(\tau)^\top \right\} dt \\ &+ A_2(\tau)\sigma_X\, dz. \end{aligned} \tag{7}$$

The expression for the drift of bond prices can be simplified by applying the expressions for A_2 and A_1 given in Appendix A:

$$\frac{dP(t)}{P(t)} = (A_2(\tau)\sigma_X \bar{\Lambda}(t) + r(t))\, dt + A_2(\tau)\sigma_X\, dz.$$

Equation (7) shows that bond prices vary with the state variables $X(t)$. The correlation between bond prices and state variables depends on the maturity of the bond through the function $A_2(\tau)$. With slight abuse of notation, we let $P(t)$ denote a vector of m bond prices, with A_2 the $m \times m$ matrix with rows equal to the corresponding values of $A_2(\tau)$.

Our framework allows for the existence of other assets besides bonds. For concreteness, we assume there exists a stock portfolio with price dynamics

$$\frac{dS(t)}{S(t)} = (\sigma_S \bar{\Lambda}(t) + r(t))\, dt + \sigma_S\, dz. \tag{8}$$

The row vector σ_S is assumed to be linearly independent of the rows of σ_X, so that the stock is not spanned by bonds. We can then group the existing assets into the vector process

$$\begin{pmatrix} dP(t) \\ dS(t) \end{pmatrix} = \operatorname{diag}\begin{pmatrix} P \\ S \end{pmatrix} (\mu(t)\, dt + \sigma\, dz), \tag{9}$$

where

$$\sigma = \begin{pmatrix} A_2\sigma_X \\ \sigma_S \end{pmatrix}, \tag{10}$$

and μ is such that

$$(\mu(t) - \iota r(t)) = \sigma\,\bar{\Lambda}(t) \tag{11}$$

with ι equal to an $(m+1) \times 1$ vector of ones. Because we have assumed there exist m nonredundant bonds, and because the stock is not redundant, the variance–covariance matrix of the assets, $\sigma\sigma^\top$ is invertible.

Equation (11) shows why this specification allows for predictable excess returns. Because $\bar{\Lambda}(t)$ is a function of the state variables $X(t)$, the instantaneous expected excess return $\mu(t) - r(t)$ is also a function of $X(t)$. The structure of $\bar{\lambda}_2$ determines how quantities that are correlated with the state variables, such as the yield spread, predict asset returns.

So far, we have described the nominal economy. Because we are interested in the strategies for an investor who cares about real wealth, it is necessary to define a process for the price level. Define a stochastic price level $\Pi(t)$ such that

$$\frac{d\Pi(t)}{\Pi(t)} = \pi(X(t), t)\,dt + \sigma_\Pi\,dz. \tag{12}$$

It is assumed that expected inflation $\pi(t)$ is affine in the state variables[7]

$$\pi(t) = \zeta_0 + \zeta X(t). \tag{13}$$

In what follows, we do not require that there exists an asset that is riskfree in real terms. In nominal terms, such an asset would have diffusion proportional to $\sigma_\Pi dz$; thus, the existence of a real riskfree asset is equivalent to the existence of a portfolio that perfectly hedges $\Pi(t)$. As long as markets are incomplete (no real riskfree asset exists), there are more sources of risk than there are independent risky assets: there are $m+1$ risky assets (m bonds and one stock), but $m+2$ sources of risk (m state variables, the stock, and the price level). As a consequence, the price of risk and the pricing kernel are not unique. Any process Λ that satisfies

$$\sigma\Lambda = \mu - r\iota \tag{14}$$

is a valid price of risk. Because (14) is a system of $m+1$ equations in $m+2$ unknowns, the solution is not unique. In what follows, $\bar{\Lambda}$ denotes the price of risk that is specified in (3), while Λ denotes a (generic) solution to (14).

Of special interest is the unique price of risk, Λ^*, that both prices and is spanned by the underlying assets. This price of risk can be found by projecting $\bar{\Lambda}$

[7] It is sufficient for the portfolio choice results to require that $r(t) - \pi(t)$ is an affine function. However, (1), or equivalently (13), is required to achieve affine bond prices.

onto the rows of σ (i.e., the loadings of asset returns on the underlying Brownian motions)

$$\Lambda^* = \sigma^\top(\sigma\sigma^\top)^{-1}\sigma\bar{\Lambda} = \sigma^\top(\sigma\sigma^\top)^{-1}\sigma\Lambda = \sigma^\top(\sigma\sigma^\top)^{-1}(\mu - r\iota). \tag{15}$$

The last two equalities hold for any price of risk satisfying (14). Because we have assumed homoscedasticity, Λ^* has the same functional form as $\bar{\Lambda}$, with

$$\lambda_1^* = \sigma^\top(\sigma\sigma^\top)^{-1}\sigma\bar{\lambda}_1 \tag{16}$$

$$\lambda_2^* = \sigma^\top(\sigma\sigma^\top)^{-1}\sigma\bar{\lambda}_2 \tag{17}$$

replacing $\bar{\lambda}_1$ and $\bar{\lambda}_2$ in (3).

The price of risk Λ^* is of interest for several reasons. First, the Cauchy inequality implies

$$\max_\sigma \frac{\sigma\Lambda^*}{\sqrt{\sigma\sigma^\top}} = \frac{(\Lambda^*)^\top\Lambda^*}{\sqrt{(\Lambda^*)^\top\Lambda^*}} = \sqrt{(\Lambda^*)^\top\Lambda^*}.$$

Thus, the norm of Λ^* equals the maximal Sharpe ratio. The maximal Sharpe ratio is always positive, even if Λ^* is not; this is because an investor can take both short and long positions in any asset. Second, any price of risk Λ can be written as a sum of Λ^* and a process that is in the null space of σ. That is,

$$\Lambda = \sigma^\top(\sigma\sigma^\top)^{-1}\sigma\Lambda + (\Lambda - \sigma^\top(\sigma\sigma^\top)^{-1}\sigma\Lambda) = \Lambda^* + \nu. \tag{18}$$

The second term, ν, satisfies $\sigma\nu = 0$, and thus is in the null space of the underlying asset returns. This term completely determines Λ. There is a one-to-one mapping between valid prices of risk Λ and processes ν in the null space of σ. We denote the pricing kernel associated with $\Lambda^*(t)$ by $\phi^*(t)$ and the pricing kernel associated with $\Lambda^*(t) + \nu(t)$ by $\phi_\nu(t)$, where

$$\frac{d\phi^*(t)}{\phi^*(t)} = -r(t)\,dt - (\Lambda^*(t))^\top\,dz, \tag{19}$$

and

$$\frac{d\phi_\nu(t)}{\phi_\nu(t)} = -r(t)\,dt - (\Lambda^*(t) + \nu(t))^\top\,dz. \tag{20}$$

While we started by defining a pricing kernel for nominal assets, we could have equivalently defined payoffs in real terms, and defined a pricing kernel for real assets. Any nominal pricing kernel $\phi_\nu(t)$ is associated with a "real" pricing kernel. For an asset with nominal value $V(s)$ at time s, the price at time t (assuming the asset pays no dividends between t and s) equals

$$V(t) = E_t\left[\frac{\phi_\nu(s)}{\phi_\nu(t)}V(s)\right]. \tag{21}$$

It follows directly from (21) that for the real payoff $V(s)/\Pi(s)$,

$$\frac{V(t)}{\Pi(t)} = E_t\left[\frac{\phi_\nu(s)\Pi(s)}{\phi_\nu(t)\Pi(t)}\left(\frac{V(s)}{\Pi(s)}\right)\right]. \tag{22}$$

Therefore, $\phi_\nu(t)\Pi(t)$ is a valid pricing kernel when asset prices are expressed in real terms. This also follows from the interpretation of $\phi_\nu(t)$ as a system of Arrow-Debreu state prices. Normalizing $\phi_\nu(0) = 1$ and $\Pi(0) = 1$, $\phi_\nu(t)$ is a ratio of units of consumption at time 0 to dollars at time t. Then $\phi_\nu(t)\Pi(t)$ is a ratio of consumption at time 0 to consumption at time t. We choose to model prices in nominal rather than real terms for ease of comparison to the affine term structure literature.

The connection between incomplete markets and the lack of a real riskfree rate can also be seen from the real pricing kernel associated with the nominal kernel $\phi_\nu(t)$. It follows from Ito's Lemma that

$$\frac{d(\phi_\nu(t)\Pi(t))}{\phi_\nu(t)\Pi(t)} = (-r(t) + \pi(t) - \sigma_\Pi(\Lambda^*(t) + \nu(t)))\,dt + (\sigma_\Pi - \Lambda^*(t) - \nu(t))\,dz. \tag{23}$$

If a real riskfree rate were to exist, its real return must equal $r(t) - \pi(t) + \sigma_\Pi(\Lambda^*(t) + \nu(t))$, the drift of the real pricing kernel. While $\pi(t)$, $r(t)$, and $\sigma_\Pi\Lambda^*(t)$ are observable (note that $\Lambda^*(t)$ can be inferred from asset prices using equation (15)), $\sigma_\Pi\nu(t)$ is not. In particular, any choice of ν satisfying $\sigma\nu = 0$ is consistent with the same asset prices, but implies different values of $\sigma_\Pi\nu$, and thus different real riskfree rates.

To summarize, the investor has access to an asset with riskless nominal return r, and $m + 1$ risky assets whose nominal price dynamics are described by (9), (10), and (11). Nominal markets are complete in that there exists a full term structure of nominal bonds.[8] However, real markets may be incomplete, because there may not exist a combination of assets spanning unexpected inflation. Equivalently, there may not exist an asset that is riskfree in real terms.

II. Optimal Portfolio Choice

In this section, we derive the optimal portfolio allocation for an investor who takes bond and stock prices as given. Section A describes the general form of the solution when there is unexpected inflation. Section B specializes to the case of an affine term structure.

A. Portfolio Choice When Inflation Cannot Be Entirely Hedged: General Results

We first solve the portfolio choice problem for an investor with power utility over terminal wealth at date T, and then generalize to the case of consumption withdrawal. We assume that the investor solves

[8] In what follows, we also consider cases where the investor has access to only a subset of the bonds (incomplete nominal markets).

$$\max_{W(T)>0} E_t \left[\frac{(W(T)/\Pi(T))^{1-\gamma}}{1-\gamma} \right], \tag{24}$$

such that $W(T)$ can be achieved by taking positions in the underlying assets with initial wealth $W(0)$:

$$\frac{dW(t)}{W(t)} = w(t)^\top (\mu(t) - r(t)\iota)\,dt + r(t)\,dt + w(t)^\top \sigma(t)\,dz, \tag{25}$$

where $w(t)$ is an $(m+1) \times 1$ vector of portfolio weights that satisfies integrability conditions. To disallow doubling strategies, we require $W(t) > 0$ for all t (see Dybvig and Huang (1988)).

To solve this problem, it is convenient to use the martingale technique of Cox and Huang (1989), Karatzas, Lehoczky, and Shreve (1987), and Pliska (1986) generalized to the case of incomplete markets by He and Pearson (1991).[9] Cox and Huang (1989) show that when markets are complete, the dynamic budget constraint (25) can be replaced by a static budget constraint analogous to the no-arbitrage condition (5) that determines bond prices. That is,

$$E[\phi(T)W(T)] = W(0), \tag{26}$$

for the unique pricing kernel $\phi(t)$. When markets are incomplete, however, wealth, like any tradeable asset, must satisfy

$$E[\phi_\nu(T)W(T)] = W(0) \tag{27}$$

for any pricing kernel ϕ_ν. In general, optimizing with respect to (27) for a particular pricing kernel produces an incorrect answer because it is not possible to replicate the resulting process for wealth by trading in the underlying assets.

The insight of He and Pearson (1991) is that it suffices to verify (27) with respect to a single pricing kernel ϕ_{ν^*}. As He and Pearson show, the incomplete-markets problem can be recast as a complete-markets problem by "adding" sufficient assets to complete the market, but setting the return process on these assets so that their weight in the investor's optimal portfolio is zero. In other words, it suffices to choose ν such that in the complete market with (unique) pricing kernel ϕ_ν, the "additional" assets are not traded by the investor. The resulting pricing kernel ϕ_{ν^*} is called the minimax kernel because it is the kernel that minimizes the investor's maximized utility; the "worst" way to add assets from the point of view of the investor is to set their return processes such that the investor does not want to trade them.

Thus, the incomplete-markets case can be solved as the complete markets case if ϕ_{ν^*}, given by

$$\frac{d\phi_{\nu^*}(t)}{\phi_{\nu^*}(t)} = -r(t)\,dt - (\Lambda^*(t) + \nu^*(t))^\top dz, \tag{28}$$

[9] Recently, Schroder and Skiadas (1999, 2002) extend this work to a broader class of stochastic processes for the state variables and to a broader class of utility functions, including recursive utility.

is used as the pricing kernel. Precisely, the investor optimizes wealth with respect to

$$E[\phi_{\nu^*}(T)W(T)] = W(0). \tag{29}$$

For some Lagrange multiplier l, the investor's first-order condition equals

$$\frac{W(T)^{-\gamma}}{\Pi(T)^{1-\gamma}} = l\,\phi_{\nu^*}(T),$$

and the optimal terminal wealth policy is given by

$$W(T) = \left(l\,\phi_{\nu^*}(T)\Pi(T)^{1-\gamma}\right)^{-\frac{1}{\gamma}}. \tag{30}$$

Substituting back into (29) gives the expression for l.[10] Given ϕ_{ν^*}, (30) describes optimal wealth. Given optimal wealth, and hence an optimal portfolio rule, ϕ_{ν^*} is determined by setting the demand for the non-traded assets to zero.

The investor's terminal wealth policy has an economic interpretation. Rearranging,

$$\frac{W(T)}{\Pi(T)} = (l\,\phi_{\nu^*}(T)\Pi(T))^{-\frac{1}{\gamma}}. \tag{31}$$

The left-hand side is equal to real wealth. The term inside parentheses on the right-hand side is proportional to $\phi_{\nu^*}(T)\Pi(T)$. This equals the real pricing kernel corresponding to the nominal kernel ϕ_{ν^*}. Thus, (31) states that the greater the price of a given state, the less the agent consumes in that state. The lower the risk aversion (γ), the more the agent adjusts terminal wealth in response to changes in the state-price density. Note, however, that ϕ_{ν^*} is also implicitly a function of γ.

The optimal portfolio allocation is derived using (30). Define a new state variable equal to the real wealth of the log utility investor if the unique price of risk were ϕ_{ν^*}. In our environment with inflation, this state variable equals

$$Z_{\nu^*}(t) = (l\,\phi_{\nu^*}(t)\Pi(t))^{-1}. \tag{32}$$

No-arbitrage implies that wealth at time t must equal the present discounted value of wealth at time T, where the discounting is accomplished by the minimax pricing kernel:

$$\begin{aligned} W(t) &= \phi_{\nu^*}(t)^{-1}E_t\left[\phi_{\nu^*}(T)\Pi(T)Z_{\nu^*}(T)^{\frac{1}{\gamma}}\right] \\ &= \Pi(t)Z_{\nu^*}(t)E_t\left[Z_{\nu^*}(T)^{\frac{1}{\gamma}-1}\right]. \end{aligned} \tag{33}$$

The next theorem characterizes the optimal wealth and portfolio weights.

[10] Solving (29) for l implies

$$l = W(0)^{-\gamma}\left(E\left(\phi_{\nu^*}(T)^{1-\frac{1}{\gamma}}\Pi(T)^{1-\frac{1}{\gamma}}\right)\right)^{\gamma}.$$

THEOREM 1: *Assume that the investor has utility over terminal wealth with coefficient of relative risk aversion* γ*. At time* t*, optimal wealth takes the form*

$$W(t) = \Pi(t) Z_{\nu^*}(t)^{\frac{1}{\gamma}} F(X(t), t, T), \tag{34}$$

where $Z_{\nu^*}(t)$ *is given by* (32)*. The minimax pricing kernel equals*

$$\frac{d\phi_{\nu^*}}{\phi_{\nu^*}} = -r\,dt - (\Lambda^* + \nu^*)^\top dz,$$

with

$$\nu^* = (1-\gamma)\big(\sigma_\Pi - \big(\sigma_\Pi \sigma^\top\big)(\sigma\sigma^\top)^{-1}\sigma\big)^\top. \tag{35}$$

The function F *satisfies the partial differential equation*

$$\begin{aligned} &\frac{1-\gamma}{\gamma}(r-\pi)F + F_X\left(K(\theta - X) + \frac{1}{\gamma}\sigma_X(\Lambda^* + \nu^*) + \frac{\gamma-1}{\gamma}\sigma_X\sigma_\Pi^\top\right) \\ &\quad + F_t + \frac{1}{2}\left(\frac{1}{\gamma}\frac{1-\gamma}{\gamma}\big((\Lambda^*+\nu^*)^\top(\Lambda^*+\nu^*) + \sigma_\Pi^\top\sigma_\Pi\big)F + \mathrm{tr}\big(F_{XX}\sigma_X\sigma_X^\top\big)\right) \\ &\quad = \left[\frac{\gamma-1}{\gamma}\right]^2 \sigma_\Pi(\Lambda^*+\nu^*)F + F_X\sigma_X(\Lambda^*+\nu^*), \end{aligned} \tag{36}$$

with boundary condition $F(X(T), T, T) = 1$.[11] *The optimal portfolio allocation equals*

$$\begin{aligned} w(t) &= \frac{1}{\gamma}(\sigma\sigma^\top)^{-1}(\mu - \iota r) + \left(1 - \frac{1}{\gamma}\right)(\sigma\sigma^\top)^{-1}\big(\sigma\sigma_\Pi^\top\big) \\ &\quad + (\sigma\sigma^\top)^{-1}\big(\sigma\sigma_X^\top\big)\frac{1}{F}(F_X)^\top. \end{aligned} \tag{37}$$

The remainder of the investor's wealth, $1 - w(t)^\top\iota$*, is invested in the nominal riskfree asset.*

The proof is given in Appendix B. The minimax price of risk equals the price of risk spanned by the existing assets plus ν^*, where ν^* equals $1-\gamma$ times the unhedgeable part of inflation risk. Thus, ν^* can be interpreted as an investor-specific measure of market incompleteness.

Equation (37) shows that the investor can be viewed as investing in $m+2$ "funds." The first fund is the portfolio that is instantaneously mean-variance efficient. It is straightforward to check that this portfolio achieves the maximum Sharpe ratio $\sqrt{(\Lambda^*)^\top\Lambda^*}$. The second fund adjusts for the fact that the first fund is mean-variance efficient in nominal rather than real terms. Together, these portfolios constitute what is known as "myopic demand," namely the optimal allocation if the investor ignores the future investment opportunity set.

It is the last term in (37) that is the focus of this study. This term represents the sum of the m hedge portfolios:

[11] The variable $\mathrm{tr}(\cdot)$ denotes the trace. The variable F_{XX} is the $m \times m$ matrix of second derivatives.

$$(\sigma\sigma^\top)^{-1}\big(\sigma\sigma_X^\top\big)\frac{1}{F}\,(F_X)^\top = \frac{1}{F}\sum_{j=1}^{M}(\sigma\sigma^\top)^{-1}\big(\sigma\sigma_{X_j}^\top\big)F_{X_j}.$$

Hedge portfolio j is formed by projecting state variable j onto the available assets. Scaling the portfolio is the sensitivity of wealth to state variables j, $\frac{1}{F}(F_{X_j})^\top$. If increases in state variable j increase wealth in the future, then the investor allocates a positive amount to the hedge portfolio $(\sigma\sigma^\top)^{-1}(\sigma\sigma_{X_j}^\top)$. Because we have assumed that there are as many nonredundant bonds as state variables, it is possible to completely hedge the state variables by trading in the underlying assets. Moreover, hedging demand for bonds is nonzero. Because bonds are the discounted value of \$1, their prices co-vary with the variables that affect the investment opportunity set, namely $X(t)$.

Also of interest is the investor's indirect utility. Cox and Huang (1989) show that it is possible to derive indirect utility from the expression for wealth. Corollary 1 generalizes this result to the case where there is unexpected inflation (and specializes to the case of power utility).

COROLLARY 1: *Define the investor's indirect utility function as follows:*

$$J(W(t), \Pi(t), X(t), t, T) = E_t\left[\frac{1}{1-\gamma}\left(\frac{W(T)}{\Pi(T)}\right)^{1-\gamma}\right]. \tag{38}$$

Then $J(W, \Pi, X, t, T)$ takes the form

$$J(W(t), \Pi(t), X(t), t, T) = \frac{1}{1-\gamma}\left(\frac{W(t)}{\Pi(t)}\right)^{1-\gamma} F(X(t), t, T)^\gamma,$$

where $F(X(t), t, T)$ is defined in Theorem 1.

The proof of Corollary 1 can be found in Appendix B.

These results generalize to the case where the investor has utility over consumption between times 0 and T. At each time, besides allocating wealth among assets, the investor also decides what proportion of wealth to consume. The investor solves

$$\max E\left[\int_0^T e^{-\rho t}\frac{(c(t)/\Pi(t))^{1-\gamma}}{1-\gamma}\,dt\right] \tag{39}$$

$$\text{s.t. } dW(t) = (w(t)^\top(\mu(t) - r(t)\iota) + r(t))W(t)\,dt + w(t)^\top\sigma W(t)\,dz - c(t)\,dt$$
$$W(T) \geq 0.$$

As shown in Wachter (2002), computing the solution to this case does not require solving a new partial differential equation.[12] As in the case of terminal wealth, the dynamic problem can be recast as static problem for an endogenous

[12] While the results in Wachter (2002) assumed that markets were complete, the same reasoning can be applied here because the adjustment for incomplete markets in the minimax pricing kernel (35) takes a particularly simple form.

pricing kernel. Using arguments similar to those in the proof of Theorem 1, it can be shown that, when the only market incompleteness comes from inflation, the investor-specific pricing kernel (ϕ_{ν^*}) for the case of intermediate consumption takes the same form as the investor-specific pricing kernel for terminal wealth. The static budget constraint is therefore equal to

$$E\left[\int_0^T c(t)\phi_{\nu^*}(t)\,dt\right] = W(0). \tag{40}$$

The following corollary describes the form of the investor's consumption policy, optimal wealth, and portfolio allocation.

COROLLARY 2: *The optimal consumption policy $c(t)$ satisfies*

$$\frac{c(t)}{\Pi(t)} = (l\,\phi_{\nu^*}(t)\Pi(t))^{-\frac{1}{\gamma}} e^{-\frac{\rho}{\gamma}t}, \tag{41}$$

where l is the Lagrange multiplier that allows (40) *to hold. Optimal wealth is given by*

$$W(t) = Z_{\nu^*}(t)^{\frac{1}{\gamma}}\,\Pi(t)\int_t^T F(X(t),t,s)e^{-\frac{\rho}{\gamma}(s-t)}\,ds, \tag{42}$$

where $Z_{\nu^*}(t)$ is defined by (32)*, and F satisfies the partial differential equation* (36)*. The optimal portfolio weights are given by* (37) *with F replaced by $\int_t^T Fe^{-\frac{\rho}{\gamma}(s-t)}$.*

Theorem 1 shows that in the homoskedastic setting of our paper, the investment opportunity set is determined by Λ^* and $r-\pi$. This can be seen from the differential equation (36), and the fact that σ_Π, σ_X, σ_S, and ν^* (by (35)) are constant. Note that r and π do not appear separately in (36), they only appear as the difference $r-\pi$. For convenience, we abuse terminology slightly and refer to $r-\pi$ as the real riskfree rate, keeping in mind that there may not exist an asset that is riskfree in real terms.[13]

B. Portfolio Allocation When the Nominal Term Structure Is Affine

Theorem 1, Corollary 1, and Corollary 2 do not require that bond yields be affine. They hold generally, as long as the investor has power utility over terminal wealth. The following corollary explicitly solves for the portfolio weights, given the assumptions on $\bar{\Lambda}$, r, and π.

[13] Indeed, the results in Section I show that this only equals the real riskfree rate if markets are completed such that the price of inflation risk is zero.

COROLLARY 3: *Assume $\bar{\Lambda}$ and $r - \pi$ are linear in the state variables $X(t)$, and that inflation and asset prices are homoscedastic, and the investor has utility over terminal wealth given by* (24). *Then F takes the form*

$$F(X(t), t, T) = \exp\left\{\frac{1}{\gamma}\left(\frac{1}{2}X(t)^\top B_3(\tau)X(t) + B_2(\tau)X(t) + B_1(\tau)\right)\right\}, \quad (43)$$

where $\tau = T - t$ and the matrix B_3, the vector B_2, and the scalar B_1 satisfy a system of ordinary differential equations. The optimal portfolio rule equals

$$\begin{aligned} w(t) = {} & \frac{1}{\gamma}(\sigma\sigma^\top)^{-1}(\mu - \iota r) + \frac{\gamma - 1}{\gamma}(\sigma\sigma^\top)^{-1}\big(\sigma\sigma_\Pi^\top\big) \\ & + \frac{1}{\gamma}(\sigma\sigma^\top)^{-1}\big(\sigma\sigma_X^\top\big)\left(\frac{B_3(\tau) + B_3(\tau)^\top}{2}X(t) + B_2(\tau)^\top\right). \end{aligned} \quad (44)$$

The remainder of the investor's wealth, $1 - w(t)^\top \iota$, is invested in the nominal riskfree asset.

The proof of Corollary 3 and the differential equations for B_3, B_2, and B_1 can be found in Appendix B. A noteworthy special case arises when risk premia are constant. Then $B_3(\tau) = 0$, as can be checked by setting $\lambda_2^* = 0$ into the differential equation (B7). The optimal portfolio allocation is constant, and F is exponential-affine. A two-factor version of this case is considered by Brennan and Xia (2002).

Why do time-varying risk premia produce a functional form for optimal wealth (43), and hence for indirect utility (by Corollary 2), that is exponential quadratic? As Campbell and Viceira (1999) discuss, the reason is that the investor can profit both when risk premia $\sigma\Lambda^*$ are especially high and positive, and when they are especially low and negative. A function for wealth that is quadratic in $X(t)$ captures this quality. Note that exponential-quadratic wealth implies a portfolio rule that is linear in the state variables.

Using Corollary 2, it is also possible to write down an explicit formula for the optimal portfolio for an investor with utility over consumption.

COROLLARY 4: *Assume $\bar{\Lambda}$ and $r - \pi$ are linear in the state variables $X(t)$, and that inflation and asset prices are homoscedastic. Suppose the investor has utility over consumption. The optimal portfolio weights equal*

$$\begin{aligned} w(t) = {} & \frac{1}{\gamma}(\sigma\sigma^\top)^{-1}(\mu - \iota r) + \frac{\gamma - 1}{\gamma}(\sigma\sigma^\top)^{-1}\big(\sigma\sigma_\Pi^\top\big) + \frac{1}{\gamma}(\sigma\sigma^\top)^{-1}\big(\sigma\sigma_X^\top\big) \\ & \times \left(\frac{\displaystyle\int_t^T F(t, t+\tau)\left(\frac{1}{2}(B_3(\tau) + B_3(\tau)^\top)X(t) + B_2(\tau)^\top\right)e^{-\frac{\rho}{\gamma}\tau}d\tau}{\displaystyle\int_t^T F(t, t+\tau)e^{-\frac{\rho}{\gamma}\tau}d\tau}\right). \end{aligned}$$

The results above show that wealth, indirect utility, and the optimal allocation are available in closed form up to the solution of ordinary differential equations. In the following sections, we estimate the parameters of the model and evaluate the implications for portfolio choice.

III. Estimation

The previous sections described optimal portfolio choice when the nominal term structure is affine and the investor has access to stock as well as bonds. In this section, we estimate a three-factor term structure model that has been shown to perform well in out-of-sample forecasting (Duffee (2002)), and in replicating the failure of the expectations hypothesis seen in the data (Dai and Singleton (2002)).[14] Our estimation differs from the estimation in these studies in that we incorporate data on equity returns, and most importantly, on inflation.[15]

There are five sources of risk in the model. The first three are due to the state variables X defined in (2), the fourth is due to the stock price S defined in (8), and the fifth is due to the price level Π defined in (12). Thus, dz is a 5×1 vector of independent Brownian motions, σ_X is a 3×5 matrix, and σ_S and σ_Π are 1×5 vectors. Without loss of generality, we order the elements of dz so that when σ_X, σ_S, and σ_Π are stacked, the resulting 5×5 matrix is lower triangular:

$$\begin{bmatrix} \sigma_X \\ \sigma_S \\ \sigma_\Pi \end{bmatrix} = \begin{bmatrix} \sigma_{X(1,1)} & 0 & 0 & 0 & 0 \\ \sigma_{X(2,1)} & \sigma_{X(2,2)} & 0 & 0 & 0 \\ \sigma_{X(3,1)} & \sigma_{X(3,2)} & \sigma_{X(3,3)} & 0 & 0 \\ \sigma_{S(1)} & \sigma_{S(2)} & \sigma_{S(3)} & \sigma_{S(4)} & 0 \\ \sigma_{\Pi(1)} & \sigma_{\Pi(2)} & \sigma_{\Pi(3)} & \sigma_{\Pi(4)} & \sigma_{\Pi(5)} \end{bmatrix}.$$

Thus, dz_1 is the risk arising from X_1, dz_2 is risk arising from X_2 that is orthogonal to the risk in X_1, dz_3 is risk arising from X_3 that is orthogonal to the risk in X_1 and X_2, etc.

In the estimation, we seek to identify

$$\Lambda^*(t) = \lambda_1^* + \lambda_2^* X(t),$$

the unique price of risk that is within the span of the underlying assets. Given the ordering for dz, it follows that λ_1^* and λ_2^* take the form

$$\lambda_1^* = \left[\lambda_{1(1)}^* \cdots \lambda_{1(4)}^* \; 0\right]^\top$$

[14] In the notation of these papers, the model we estimate is known as $A_0(3)$ because it contains three factors and no square root processes.

[15] There is a substantial literature on using yields on nominal bonds to extract expected inflation. This includes Boudoukh, Richardson, and Whitelaw (1994), Fama (1975), Fama and Gibbons (1982), and Mishkin (1981) who use a regression approach, as well as Ang and Bekaert (2003), Boudoukh (1993), Pennacchi (1991), and Sun (1992), who estimate expected inflation within a term structure framework that precludes the existence of arbitrage.

and

$$\lambda_2^* = \begin{pmatrix} \lambda_{2(1,1)}^* & \cdots & \lambda_{2(1,3)}^* \\ \vdots & & \vdots \\ \lambda_{2(4,1)}^* & \cdots & \lambda_{2(4,3)}^* \\ 0 & 0 & 0 \end{pmatrix}.$$

The variables λ_1^* and λ_2^* have zeros in the fifth row because both bonds and stocks load only on the first four Brownian motions. Otherwise, λ_1^* and λ_2^* would not be within the span of σ as required.[16]

As Dai and Singleton (2000) discuss, the processes for X, Λ^*, and r have too many degrees of freedom to be identified by the data. For example, it is not possible to simultaneously identify θ and δ_0. Following Dai and Singleton (2000) and Duffee (2002), we set $\theta = 0_{3\times1}$, and specify that K is lower triangular. Further, we set

$$\sigma_X = [I_{3\times3} \quad 0_{3\times2}] \tag{45}$$

analogously to Dai and Singleton and Duffee who set σ_X equal to the identity matrix.

With the restrictions described above, all of the parameters in the model can, in principle, be identified. In practice, the large number of parameters in such models has led to concerns of over-fitting. We follow Duffee (2002) in further restricting the matrix K and the price of risk λ_2^* in order that the estimation be more reliable. Given the form of σ_X (and because bonds load only on the state variables), the first three rows of λ_1^* and λ_2^* are determined by risk premia on bonds and can be identified from term structure data. We place the same restrictions on these elements of λ_2^* as does Duffee (2002). In addition, we restrict the fourth row of λ_2^* so that the equity premium is constant. We set this requirement because of the difficulty in identifying three separate sources of variation in the equity premium that all arise from the term structure, and because the focus of this paper is on bond return, rather than stock return, predictability. Because σ_S is determined from the variance–covariance matrix of bond and stock returns, and because the first three and fifth rows of λ_1^* and λ_2^* are determined, the equation for the equity premium is given by

$$\sigma_S\left(\lambda_1^* + \lambda_2^* X(t)\right) = \eta_0.$$

Note that this is a system of four equations in four unknowns. The fourth element in λ_1^* is determined by

$$\sigma_S \lambda_1^* = \eta_0, \tag{46}$$

while the three elements in the fourth row of λ_2^* are determined by

$$\sigma_S \lambda_2^* = 0_{1\times3}. \tag{47}$$

[16] In what follows, we also consider the case of incomplete nominal markets. For these cases, Λ^* must be adjusted further so that it is within the span of the existing assets.

Table I

Processes for the Riskfree Rate and Expected Inflation

The three-factor model described in Section III is estimated using monthly data on bond yields, inflation, and stock returns from 1952 to 1998. The nominal interest rate $r(t) = \delta_0 + \delta_1 X_1(t) + \delta_2 X_2(t) + \delta_3 X_3(t)$, expected inflation equals $\pi(t) = \zeta_0 + \zeta_1 X_1(t) + \zeta_2 X_2(t) + \zeta_3 X_3(t)$. The process for X is given by $dX(t) = -KX(t)\,dt + \sigma_X dz(t)$, where σ_X is shown in Table II. Outer product standard errors are given in parentheses. Parameter values are annual and in natural units.

Panel A: Constant Terms

δ_0	ζ_0
0.056	0.040
(0.034)	(0.026)

Panel B: Coefficients on State Variables

	X_1	X_2	X_3
ζ_i	0.018	0.018	0.007
	(0.002)	(0.004)	(0.0005)
δ_i	0.018	0.007	0.010
	(0.0003)	(0.0009)	(0.0003)
$K_{1,i}$	0.576	0	0
	(0.027)		
$K_{2,i}$	0	3.343	0
		(0.210)	
$K_{3,i}$	−0.421	0	0.083
	(0.170)		(0.055)

Rather than estimate the fourth row of λ_1^* and λ_2^* directly, we estimate η_0 and back out λ_1^* and λ_2^* using (46) and (47), respectively.

Our bond data consist of monthly observations on zero coupon yields for 3-month, 6-month, and 1-, 2-, 5-, and 10-year U.S. government bonds. The bond data are available from the web site of Gregory Duffee. Monthly observations on the CPI and on returns on a broad stock index are available from CRSP. The sample begins in 1952 and ends in 1998. Following Duffee (2002), we assume that prices on the 3-month, 1-year, and 5-year bonds are measured with normally distributed errors. The model implies that state variables, stock returns, and realized inflation are jointly normally distributed. The parameters are thus δ_0, δ, ζ_0, ζ, K, λ_1^*, λ_2^*, σ_S, σ_Π, and η_0, and the variance–covariance matrix of the errors. We estimate the model using maximum likelihood, an alternative to the Generalized Method of Moments approach of Gibbons and Ramaswamy (1993). Details are contained in Appendix E.

Tables I–III describe the results from our estimation. Because the yields are in annual terms, time is in years. As shown in Table I, the parameters δ_0 and ζ_0 equal 0.056 and 0.040, respectively. The value of δ_0 is approximately equal to the mean of the 3-month Treasury bill return in the data, 0.055. The value of ζ_0 is approximately equal to the mean of inflation in the data, which is 0.039. While the ability to match these values may seem like a natural property for

Table II

Volatility Matrix

Estimates of loadings on the Brownian motions. For example, the unpredictable component of the stock price is given by $\sigma_S(1)\,dz_1 + \sigma_S(2)\,dz_2 + \sigma_S(3)\,dz_3 + \sigma_S(4)\,dz_4 + \sigma_S(5)\,dz_5$. The entries of σ_X and the last entry of σ_S cannot be identified from the data; they are set equal to the values below without loss of generality. Outer product standard errors are in parentheses.

	dz_1	dz_2	dz_3	dz_4	dz_5
	1	0	0	0	0
σ_X	0	1	0	0	0
	0	0	1	0	0
σ_S ($\times 100$)	−1.255	0.572	−2.946	14.277	0
	(0.578)	(0.592)	(0.632)	(0.304)	–
σ_Π ($\times 100$)	0.001	−0.011	0.133	−0.084	0.911
	(0.042)	(0.042)	(0.042)	(0.042)	(0.024)

Table III

Prices of Risk

Estimates of the price of risk $\Lambda^* = \lambda_1^* + \lambda_2^* X(t)$ and of the equity premium. The ith Brownian motion is denoted by $z_i(t)$. The first three rows of λ_1^* and λ_2^* control risk premia on bonds (because σ_X takes the form shown in Table II). The fourth row is determined by the equity premium $\sigma_S \Lambda^* = \eta_0$. The fifth row is zero by construction. Outer product standard errors are in parentheses. Parameter values are annual and in natural units.

Equity Premium η_0	0.075			
	(0.025)		λ_2^*	
Source of Risk	λ_1^*	X_1	X_2	X_3
$z_1(t)$	−0.563	0	1.754	0
	(0.212)		(0.078)	
$z_2(t)$	−0.245	0	−1.815	0
	(0.078)		(0.169)	
$z_3(t)$	−0.219	0.537	0.376	−0.082
	(0.052)	(0.174)	(0.117)	(0.055)
$z_4(t)$	0.440	0.111	0.305	−0.017
$z_t(5)$	0	0	0	0

a model to have, as Campbell and Viceira (2001) discuss, it is not guaranteed that models such as this one fit time series means. In fact, the affine models investigated by Duffee (2002) all result in an estimate of δ_0 that is too low compared to the mean of the short-term interest rate.[17] Surprisingly, including inflation in the estimation helps to estimate this parameter.

Table II describes the elements of σ_S and σ_Π. The first and third elements of σ_S are negative, consistent with a positive correlation between bond and stock

[17] Duffee ends his sample in 1994. This does not account for the difference, however. We estimate the $A_0(3)$ model without inflation and find $\delta_0 = 0.044\%$, even when we include the last 4 years of the sample.

Table IV

Asset Return Correlations

The first panel shows conditional correlations of asset returns implied by the parameter values in Tables I to III. For example, the conditional correlation between returns on the 1-year bond and on the stock is equal to $A_2(1)\sigma_X\sigma_S^\top(A_2(1)\sigma_S\sigma_S^\top A_2(1)^\top)^{-1/2}(\sigma_S\sigma_S^\top)^{-1/2}$ by (7). The second panel shows unconditional correlations from the data.

	1-Year Bond	5-Year Bond	10-Year Bond	Stock
		Panel A: Model		
1-year	1.000	0.878	0.741	0.191
5-year		1.000	0.950	0.208
10-year			1.000	0.212
Stock				1.000
		Panel B: Data		
1-year	1.000	0.853	0.734	0.190
5-year		1.000	0.932	0.192
10-year			1.000	0.214
Stock				1.000

returns (indeed, as Table IV shows, the correlation is positive). The fourth element of σ_S (by far the largest) represents the component of stock returns orthogonal to bond returns. The fourth element of σ_Π is negative and significant, consistent with a negative correlation between unexpected changes in the price level and stock returns. Moreover, note that the correlation between unexpected inflation and the first and the third state variables is positive, while the correlation between unexpected inflation and the second state variable is negative. Because bond returns are negatively correlated with the first and third state variables (Table V), but positively correlated with the second state variable, this is consistent with a negative correlation between unexpected inflation and bond returns. The estimates in Table II imply that the volatility of unexpected inflation, $\sqrt{\sigma_\Pi\sigma_\Pi^\top} = 0.93\%$ per annum. This is close to, but smaller than the volatility of realized inflation in the data (1.17%). This makes sense; the state variables add information and thus reduce the volatility.

Other than δ_0 described above, the parameters that we estimate for the term structure are very close to those found by Duffee (2002).[18] As Table III shows, the restrictions on λ_2^* imposed above imply that two factors determine time-varying risk premia on bonds. The first is given by the transitory factor X_2, while the second is a linear combination of X_1, X_2, and X_3, and hence is more persistent. Table III also shows that the estimated equity premium equals 7.5%.[19]

[18] The variance–covariance matrix for the errors, which we do not report, is nearly identical to that found by Duffee (2002).

[19] We find that the sample equity premium $\frac{12}{563}\sum_{t=1}^{563}[\log S_{t+1} - \log S_t - \frac{y(t,t+0.25)}{12}] = 6.25\%$. This value differs from our estimate of η_0 in part because the maximum likelihood estimate of δ_0 is not exactly equal to $\frac{1}{563}\sum_{t=1}^{563} y(t, t+0.25)$, and in part because this sample mean ignores Jensen's inequality, which η_0 takes into account (see, e.g., (E2)).

Table V

Correlations between Asset Returns, Innovations to Inflation, and Innovations to the Investment Opportunity Set

Panel A shows conditional correlations between asset returns and innovations to inflation (Π); expected inflation (π); the nominal interest rate (r); and the "real" interest rate ($r - \pi$) implied by the parameter values in Tables I–III. Panel B shows conditional correlations between asset returns and innovations to risk premia on the 1-year, 5-year, and 10-year bonds. Panel C shows conditional correlations between asset returns and innovations to the state variables. For example, the correlation between returns on the 1-year bond and innovations to Π is equal to $A_2(1)\sigma_X\sigma_\Pi^\top(A_2(1)\sigma_\Pi\sigma_\Pi^\top A_2(1)^\top)^{-1/2}(\sigma_\Pi\sigma_\Pi^\top)^{-1/2}$ by (7).

	1-Year Bond	5-Year Bond	10-Year Bond	Stock
		Panel A: Interest Rate and Inflation		
Π	−0.085	−0.118	−0.137	−0.117
π	−0.459	−0.112	−0.117	−0.087
r	−0.779	−0.478	−0.418	−0.148
$r - \pi$	−0.477	−0.704	−0.570	−0.091
		Panel B: Risk Premia		
$\mu_1 - r$	−0.213	−0.471	−0.341	−0.031
$\mu_5 - r$	−0.157	−0.445	−0.336	−0.026
$\mu_{10} - r$	−0.034	−0.380	−0.319	−0.015
		Panel C: State Variables		
X_1	−0.764	−0.366	−0.143	−0.086
X_2	0.313	0.512	0.345	0.039
X_3	−0.564	−0.777	−0.927	−0.201

Figures 1 to 3 illustrate the implications of the model for average yield spreads, standard deviations of yield spreads, and Campbell–Shiller long-rate regressions. Each figure plots the values in the data ("sample") and the values implied by the model. Following Dai and Singleton (2002), we construct 95% confidence bands by simulating 500 sample paths from our model with length equal to the sample path in the data. Figures 1 and 2 show that the model implies average yield spreads and standard deviations of yield spreads close to those found in the data. The confidence bands reflect the well-known result that means are estimated much more imprecisely than variances. In both cases, the data fall well within the error bands implied by the model. We conclude that the model does a reasonable job of fitting the cross-sectional moments of bond yields. Because the model must fit cross-sectional and time-series moments together, the fit to the cross section is not automatic.

Because our aim is to study the implications of the expectations puzzle for investors, it is especially important to determine whether the model accounts for the expectations puzzle found in the data. To do so, we follow the approach of Dai and Singleton (2002) and check whether the model replicates the empirical findings of Campbell and Shiller (1991). Dai and Singleton explain the connection

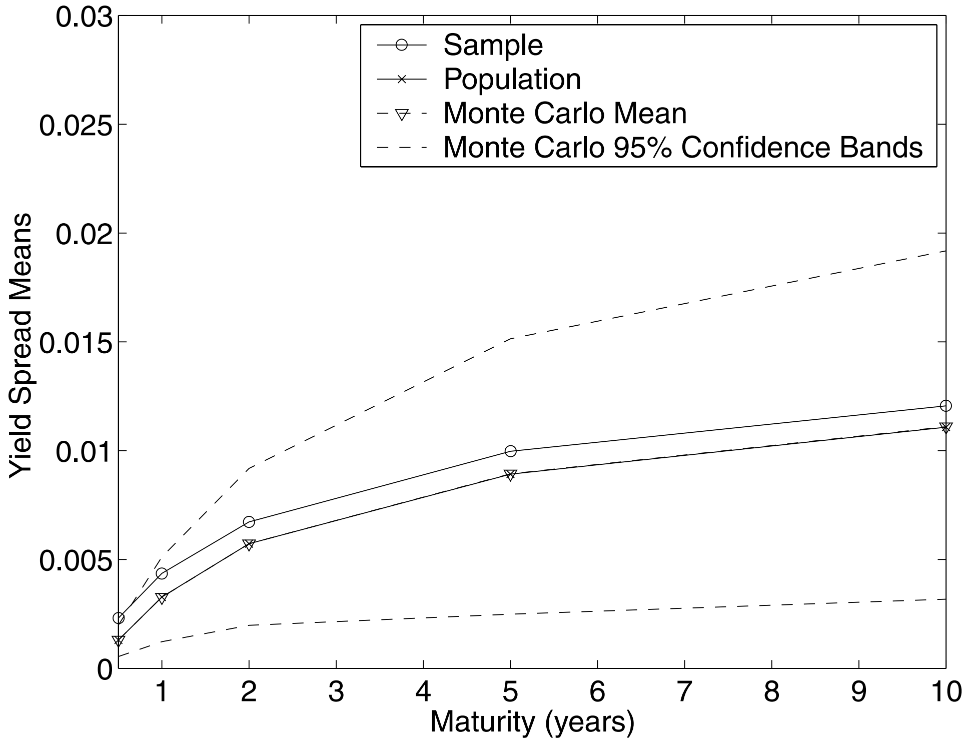

Figure 1. Model-implied yield spread means. Bond yields are in annual terms, and defined as in equation (6). The short-term yield has maturity of 3 months. Sample means are calculated using data from 1953 to 1998 on bonds of selected maturities. The population means and the means from the Monte Carlo experiment lie on top of each other.

between the Campbell–Shiller regressions and time-variation in risk premia in detail.

Figure 3 plots the slope coefficients from regressions of quarterly changes in yields on the scaled yield spread, as described in Campbell and Shiller (1991). If the expectations hypothesis held, the coefficients would be identically equal to 1. Instead, Campbell and Shiller find coefficients that are negative and decrease with maturity.[20] Figure 3 replicates this result in our data and shows that the model captures both the negative coefficients and the downward slope. Except for values at the very short end of the term structure, the data fall within the 95% confidence bands implied by the model. It is apparent from Figure 3 that the model captures the failure of the expectations hypothesis found in the data. To the extent that the failure of the expectations hypothesis is a bit less extreme

[20] Backus et al. (2001) note that the regression coefficients are noisier at longer maturities, and in fact when deviations from the expectations hypothesis are measured using forward-rate regressions, bonds at longer maturities appear to come closer to satisfying the expectations hypothesis than bonds at shorter maturities.

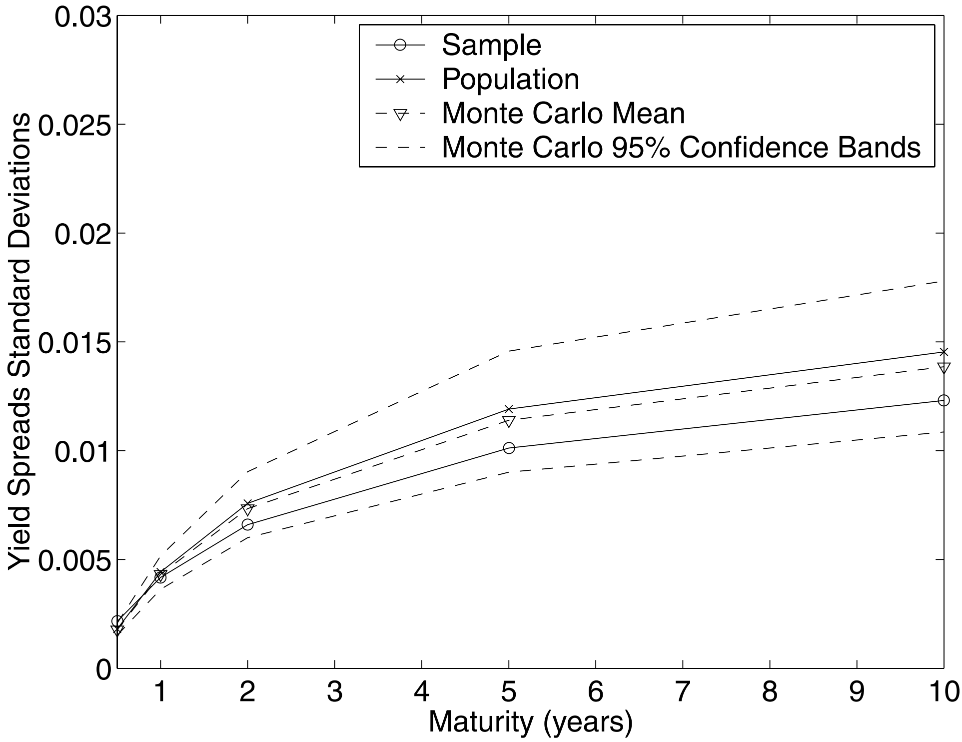

Figure 2. Model-implied yield spread standard deviations. Yields are in annual terms, and defined as in equation (6). The short-term yield has maturity of 3 months. "Sample" refers to yield spreads calculated using data from 1953 to 1998 on bonds of selected maturities.

in the model than the data, we may understate the implications for long-run investors.[21]

Figure 4 plots the time series of monthly realized inflation, and our expected inflation series constructed from the state variables using the relationship

$$\pi(t) = \zeta_0 + \zeta X(t),$$

where values for ζ_0 and ζ come from the maximum likelihood estimation described above, and are given in Table I. Our joint estimation procedure allows inflation to influence the dynamics of state variables. In practice, however, this

[21] Expectations hypothesis regressions are subject to small-sample biases that could go in either direction (Bekaert, Hodrick, and Marshall (1997, 2001), Stambaugh (1999), Valkanov (1998)). Longstaff (2000b) finds that tests fail to reject the expectations hypothesis at the short end of the term structure and argues that the failure of the expectations hypothesis may be due to a liquidity premium in Treasury Bill rates. Bekaert and Hodrick (2001) argue that standard tests tend to reject the expectations hypothesis even when it is true. They find, however, that the data remain inconsistent with the expectations hypothesis, even after adjusting for small-sample properties. Accounting for these biases within the investment decision is beyond the scope of this manuscript, but will be pursued in future work.

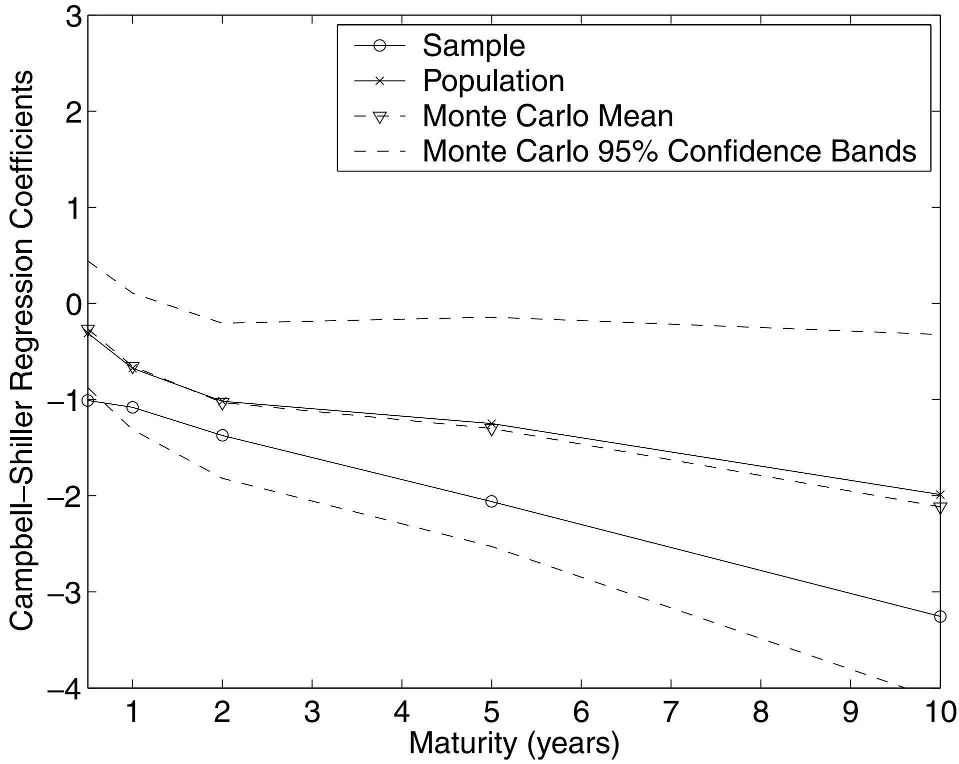

Figure 3. Model-implied coefficients on Campbell and Shiller (1991) long-rate regressions. Quarterly changes in yields, $y(t,s) - y(t+\frac{1}{4}, s)$, are regressed on the spread between the $(s-t)$-year bond, and the 3-month bond, scaled by $1/(4(s-t)-1)$. "Sample" refers to yield spreads calculated using data from 1953 to 1998 on bonds of selected maturities.

effect is small, and except for the effect on δ_0 described above, our parameter values are close to what we would find by first estimating the term structure model, and then regressing realized inflation on the factors. This latter strategy would, of course, understate the standard errors on ζ.

Figure 4 shows that our expected inflation series does indeed forecast realized inflation. In fact, expected inflation accounts for 37% of the variance of realized inflation. It is worth emphasizing that these results come about even though the factors $X(t)$ are linear combinations of yields alone. Thus, long-term bond yields contain substantial information about future inflation.

Figure 5 plots the time series for the nominal interest rate $r(t)$ implied by the model. While not shown in the graph, $r(t)$ is essentially equal to the 3-month yield. The difference between the nominal interest rate $r(t)$ and $\pi(t)$, which we informally refer to as the real interest rate, is also shown on the graph. This series is positive through nearly the entire sample. Thus, the expected inflation and real riskfree rate implied by the model have reasonable time-series properties.

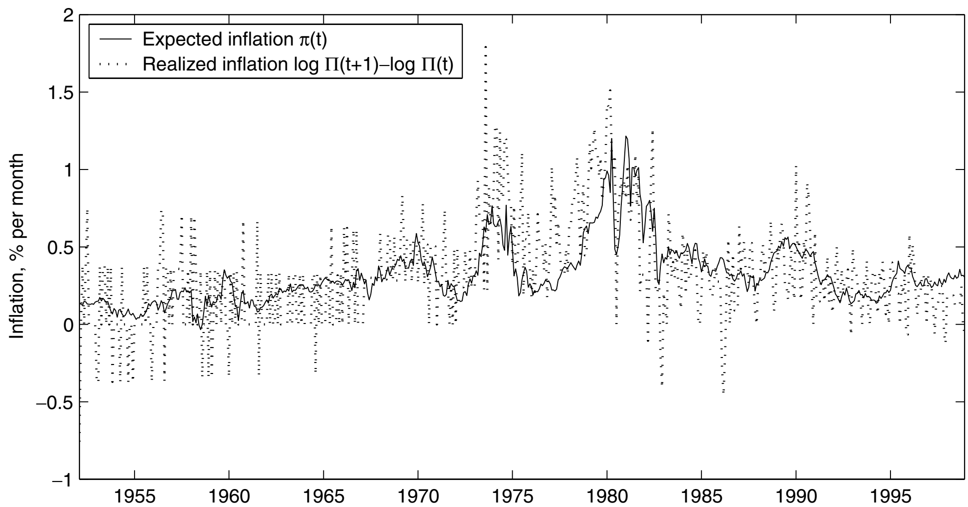

Figure 4. Realized and expected inflation. Dotted lines denote realized inflation (from data); solid lines denote expected inflation $\pi(t)$ implied by the model and parameter estimates in Table I.

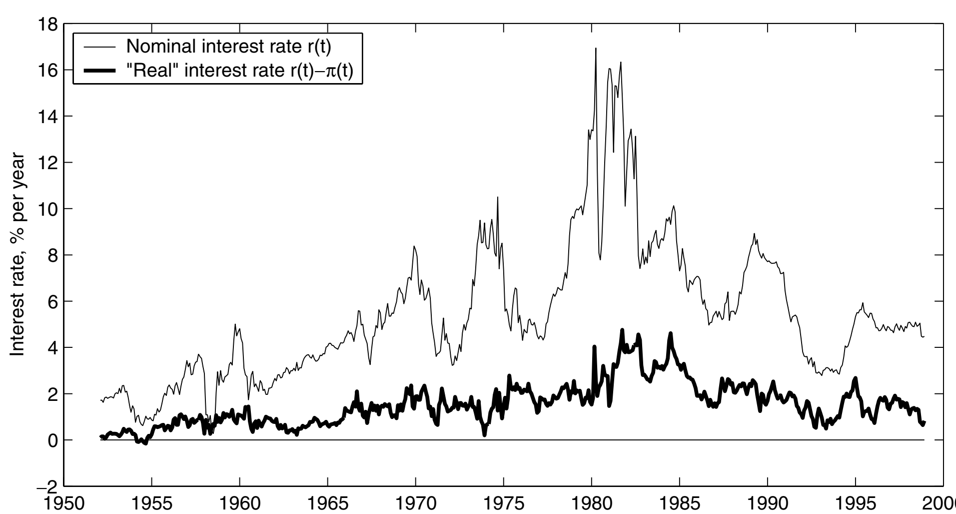

Figure 5. The nominal and real riskfree rate. The thin line denotes the nominal riskfree rate implied by the model and parameter estimates in Table I. The thick line denotes the real riskfree rate (the nominal rate minus expected inflation) implied by the model and parameter estimates in Table I.

The results in Section II show that the real interest rate $r - \pi$ and the price of risk Λ^* are the important quantities for investors. The top panel of Figure 6 plots the time series of risk premia (a linear transformation of Λ^*) for the 1-, 5-, and 10-year bonds implied by the model. As Figure 6 shows, risk premia

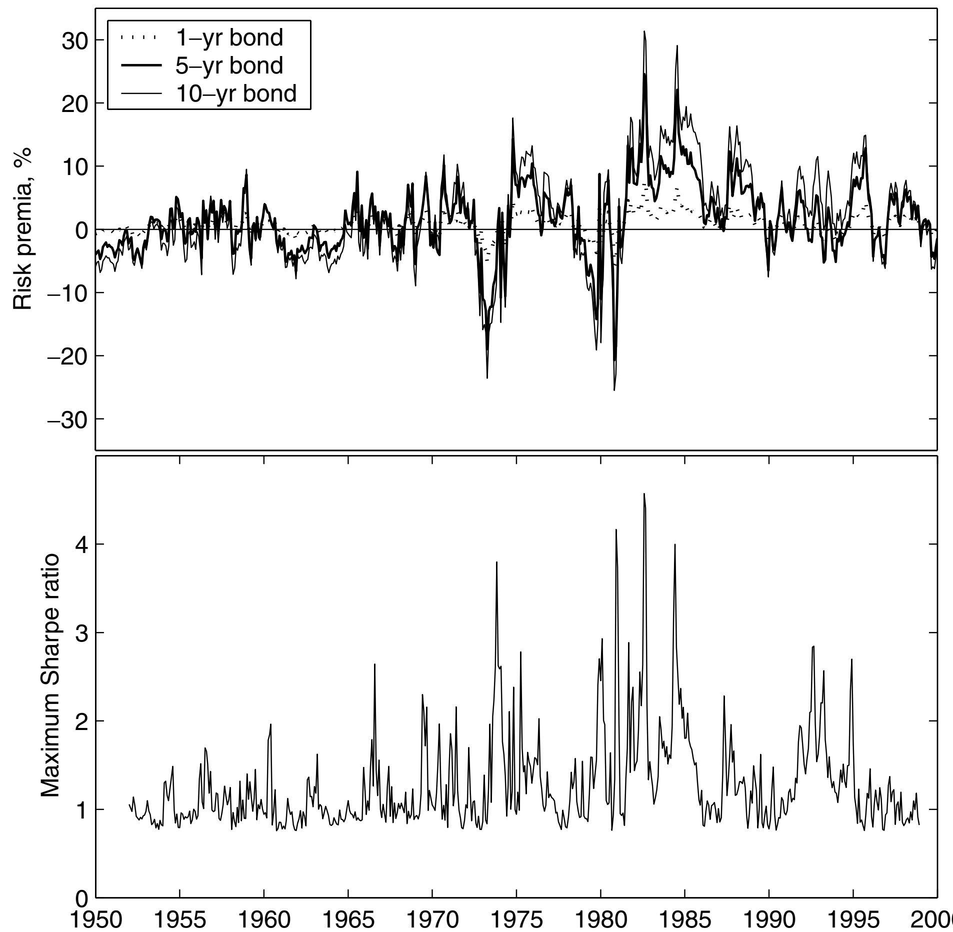

Figure 6. Risk premia and the maximum sharpe ratio. The top panel shows risk premia (in annual percentages) on long-term bonds implied by the model. The risk premium is defined as the instantaneous expected excess return over the nominal riskfree rate. The bottom panel shows the time series of the maximum Sharpe ratio achievable by taking a position in bonds, stock, and the nominal riskfree asset.

are highly volatile, especially in the latter half of the sample. Table III implies that there are two factors driving risk premia: the first is the highly transitory second state variable, the second is a linear combination of all three state variables that is much more persistent. Nonetheless, all three risk premia appear to move closely together. This is consistent with the findings of Cochrane and Piazzesi (2002), who show that a single factor can explain much of the time-variation in expected excess returns on bonds. The bottom panel of Figure 6 plots the time series of the maximal Sharpe ratio $\sqrt{(\Lambda^*)^\top \Lambda^*}$. The Sharpe ratio tends to take on its largest values when excess returns are unusually negative or positive. This large variation in investment opportunities suggests that the optimal allocation to bonds will also vary substantially as a function of the state variables.

Taking the results in this section together, we conclude that our model succeeds in capturing important features of the term structure and of inflation. The next section considers the implications of our parameter estimates for portfolio choice.

IV. Portfolio Allocation under the Failure of the Expectations Hypothesis

This section combines the theoretical results from Section II with the parameter estimates from Section III to evaluate the implications of the failure of the expectations hypothesis for long-horizon investors. The failure of the expectations hypothesis could affect the optimal portfolio in two ways. First, the myopic portfolio, $\frac{1}{\gamma}(\sigma\sigma^{\top})^{-1}(\mu - \iota r) + \frac{\gamma-1}{\gamma}(\sigma\sigma^{\top})^{-1}(\sigma\sigma_{\Pi}^{\top})$, depends directly on risk premia. If risk premia vary, so does myopic demand. Second, time-varying risk premia imply that investment opportunities vary over time (as long as changes in risk premia are not directly offset by changes in volatility). As Merton (1971) shows, the investor hedges these changes in the investment opportunity set, implying that the optimal allocation is not mean-variance efficient. Hedging demand causes the optimal portfolio for a long-horizon investor to differ from the optimal portfolio for an investor with a short horizon. Both effects are present in theory. The question is, are they economically significant?

A. *Optimal Allocation between a Long-Term Bond, Stock, and the Nominal Riskfree Asset*

To investigate the effect of time-varying risk premia on optimal portfolios, we first consider the case where the investor has access to a single long-term bond, stock, and a nominally riskfree asset. This case allows us to temporarily abstract from questions pertaining to the optimal composition of the bond portfolio and focus on the horizon properties, taking the composition as given.[22] The results in Theorem 1 apply only to the case where nominal markets are complete, namely when there are the same number of long-term bonds as state variables. However, they are easily modified for the case of incomplete nominal markets. Optimal wealth and allocation to long-term bonds still take the same form as in Corollary 3. Theorem 1 and Corollary 3 are extended to the incomplete-market case in Appendix C.[23]

Figure 7 plots the optimal allocation for the investor who allocates wealth between a 5-year bond, stock, and the nominally riskfree asset. The investor

[22] Note, however, that there is a cost to restricting the maturity of the bond. Brennan and Xia (2002) show in a related setting that, if the investor is allowed to trade only one bond, the optimal maturity of the bond depends on the investor's horizon.

[23] The results for utility over consumption (Corollary 4) have no straightforward extension.

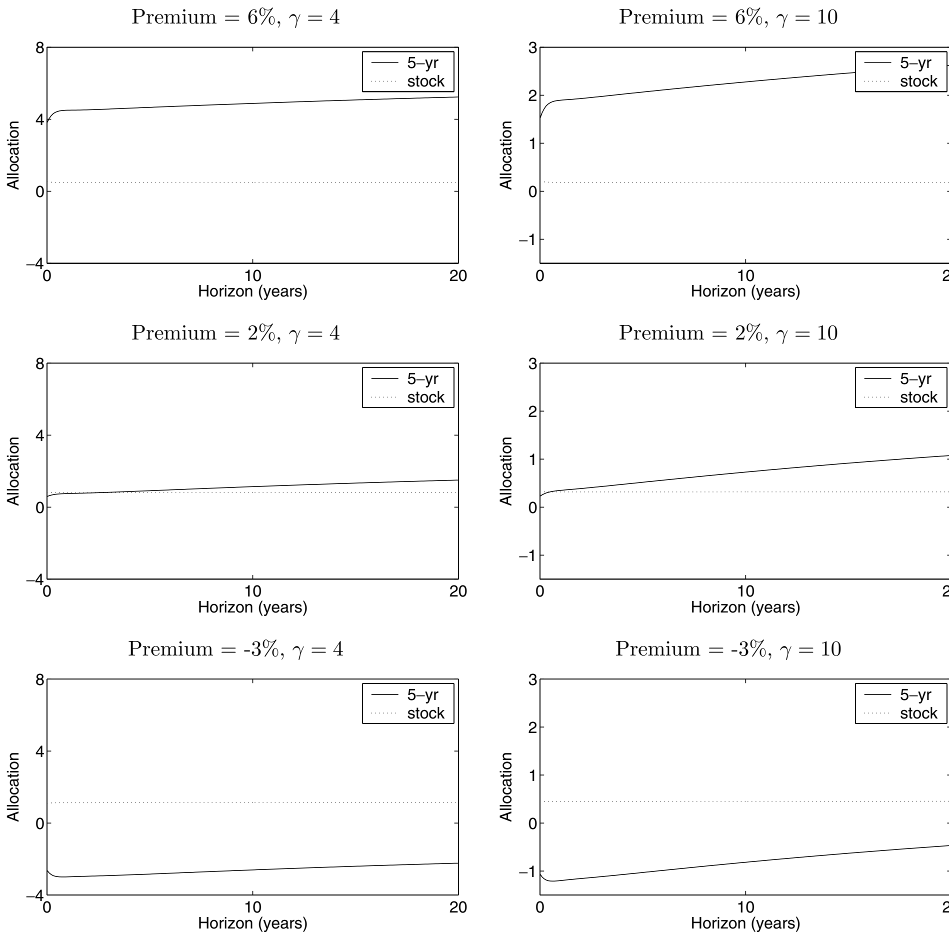

Figure 7. Optimal allocation as a function of horizon for an investor with access to a 5-year bond, the stock market, and the nominal riskfree asset. Shown are the allocations to the bond and the stock; the allocation to the riskfree asset is one minus the sum of these quantities. The state variables X_2 and X_3 are set equal to zero, while X_1 is set equal to minus one standard deviation (1.9) in the top panel, 0 in the middle panel, and plus one standard deviation in the bottom panel. "Premium" refers to the risk premium on the 5-year bond implied by each value of $X = (X_1, X_2, X_3)$. Risk aversion $\gamma = 4$ (left panel) or 10 (right panel).

is assumed to have utility over wealth at the end of the horizon. In the left panel, risk aversion $\gamma = 4$, in the right panel $\gamma = 10$. Both myopic demand and hedging demand depend on the current premia on bonds over the riskfree rate. Thus, the optimal allocation is a function of the state as well as the horizon. In order to understand how the optimal portfolio varies with the state, we plot the optimal allocation when the state variables are equal to their long-run mean of zero, and then we vary each state variable by two unconditional standard deviations. The results are similar in each case, so we discuss only the effects of varying X_1.[24]

The risk premium on the 5-year bond equals 2% per annum when the state variables are at their long-run mean, 6% when X_1 is two standard deviations below its long-run mean, and -3% when X_1 is two standard deviations above its long-run mean. The negative relationship between the risk premium on the 5-year bond and X_1 is implied by the parameter estimates in Table III and the correlation between the 5-year bond and the state variables in Table V. In particular, Table III shows that $\lambda_{2(3,1)} > 0$ and $\lambda_{2(i,1)} = 0$ for $i = 1, 2$. This means that the price of risk associated with the third Brownian motion is increasing and the price of risk for the first two Brownian motions is constant in X_1. Because state variables can be exactly identified with Brownian motions, the price of X_3 risk is increasing in X_1. Because bond prices load negatively on X_3, risk premia on bonds are decreasing in X_1. Note that the equity premium is constant in X_1 by construction. The fourth row of λ_2^* is set so that $\sigma_S \lambda_2^* = 0$.

The top panel in Figure 7 shows the optimal allocation when X_1 is two standard deviations below its long-run mean (risk premium = 6%), the middle panel shows the allocation when X_1 is at its long-run mean (risk premium = 2%), and the bottom panel shows the allocation when X_1 is two standard deviations above its long-run mean (risk premium = -3%). The myopic allocation is equal to the y-intercept, because, under power utility, it is independent of horizon. Not surprisingly, the lower is X_1 (and the greater are risk premia), the greater is the myopic allocation to the 5-year bond. When risk premia are at their long-run mean, the optimal allocation to bonds is positive. When risk premia are negative, the optimal allocation involves taking a short position in the long-term bond. Because of the positive correlation between returns on the 5-year bond and on the stock, an increase in the risk premium on the 5-year bond leads to a lower allocation to the stock. A comparison between the panels shows that the investor with risk aversion $\gamma = 4$ times the market more aggressively than the investor with risk aversion equal to 10.

There are strong horizon effects for long-term bonds.[25] For values of X_1 implying positive bond premia, the allocation to the 5-year bond rises steadily with the horizon. When bond premia are negative, the optimal allocation initially falls, but then rises after a horizon of about 1 year. The difference between short-horizon and long-horizon investors is economically large. For example, when X_1 is at its long-run mean, the myopic investor with $\gamma = 10$ allocates 20% of his wealth to the long-term bond. An investor with $\gamma = 10$ and a horizon of 20 years, by contrast, allocates over 100% of his wealth to the long-term bond.

[24] The unconditional variance–covariance matrix of the state variables can be calculated using the results of Appendix E. The unconditional standard deviation is 0.93 for X_1, 0.39 for X_2, and 3.0 for X_3. Varying X_3 has smaller effects on myopic demand, which can be seen by comparing $0.93\lambda_2^*(3,1)$ to $3\lambda_2^*(3,3)$. Because X_3 is a more persistent variable, the effects on hedging demand are larger. By contrast, X_2 has a larger effect on risk premia, and thus on myopic demand. However, its effects on hedging demand are smaller because it is much less persistent.

[25] Relative to the long-term bond, stocks have a low correlation with the investment opportunity set and, thus, negligible horizon effects.

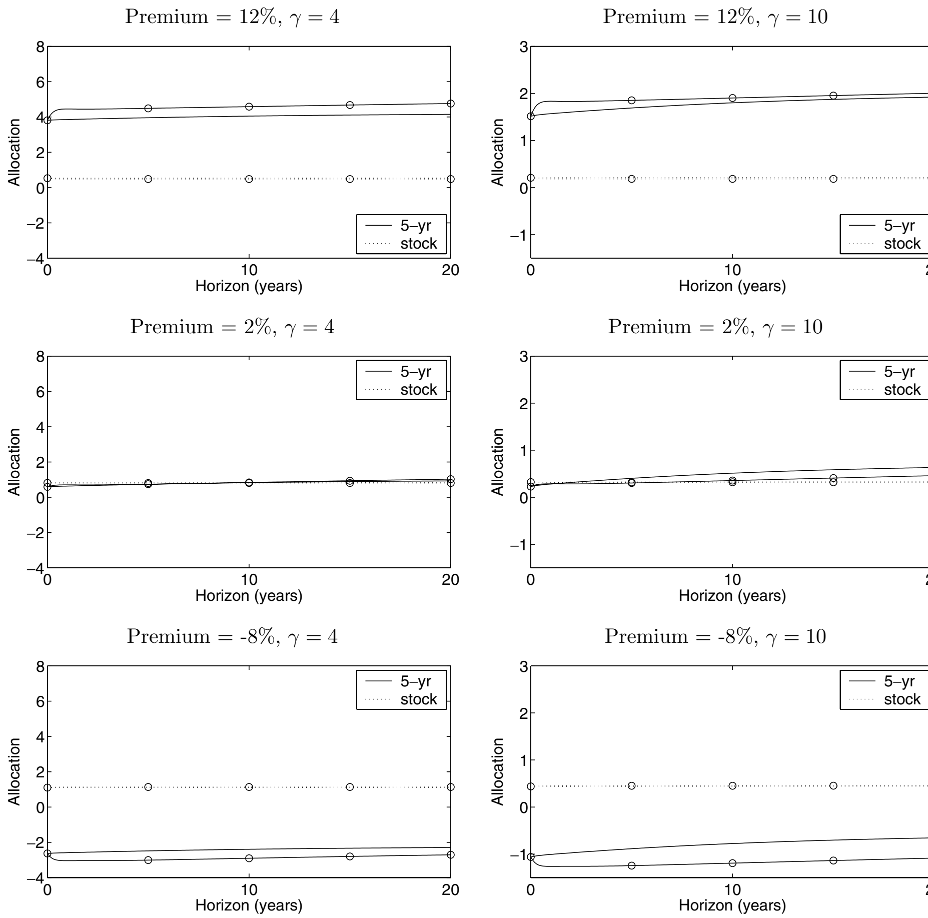

Figure 8. Allocation when the investor hedges only the real riskfree rate (plain lines) and when the investor hedges only risk premia (lines with circles) as a function of horizon. The investor has utility over terminal wealth and access to a 5-year bond, the stock market, and the nominal riskfree asset. Shown are the allocations to the bond and the stock; the allocation to the riskfree asset is one minus the sum of these quantities. For the stock, the two allocations lie on top of each other. The state variablels X_2 and X_3 are set equal to zero, while X_1 is set equal to minus one standard deviation (1.9) in the top panel, 0 in the middle panel, and plus one standard deviation in the bottom panel. "Premium" refers to the risk premium on the 10-year bond implied by each value of $X = (X_1, X_2, X_3)$. Risk aversion $\gamma = 4$ (left panel) or 10 (right panel).

While hedging demand for $\gamma = 4$ is smaller than that for $\gamma = 10$ as a proportion of myopic demand, it is still economically large. When the state variables are at their long-run mean, hedging demand more than doubles the allocation to the 5-year bond for the $\gamma = 4$ investor.

What drives the horizon effects seen in Figure 7? As discussed in Section II, hedging demand arises from two sources. One is time-variation in risk premia, the other is time-variation in the real riskfree rate, $r - \pi$. To separate out these two effects, Figure 8 plots the allocation when the investor has the correct myopic demand but she hedges only time variation in the real riskfree rate (plain lines), and the allocation when the investor has the correct myopic demand but she hedges only time variation in risk premia (lines with circles).

Suppose first that the investor hedges only time variation in the real riskfree rate (i.e., in the calculation of hedging demand, risk premia are assumed to be constant). This allocation is given by (44), the same equation that defines the optimal allocation, but with λ_2^* set equal to zero in the equations for $B_3(\tau)$, $B_2(\tau)$, and $B_1(\tau)$. Figure 8 shows that hedging demand induced by time-variation in the real riskfree rate is positive and increasing in the horizon, though it is substantially smaller than the full hedging demand shown in Figure 7.

Hedging demand resulting from the real riskfree rate is positive because long-term bond prices and the real riskfree rate are negatively correlated (Table V). A multi-period investor chooses the optimal portfolio not only to maximize the Sharpe ratio, but also so that future wealth has the "right" correlation with future investment opportunities. For $\gamma > 1$, the investor has lower marginal utility of wealth when the real riskfree rate is high; the income effect dominates (a higher riskfree rate makes him richer, he can afford a lower payoff in those states). For $\gamma < 1$, the investor has lower marginal utility of wealth when the riskfree rate is low; the substitution effect dominates (wealth is more valuable when the riskfree rate is higher because it can be invested at a higher rate). Thus, the investor with $\gamma > 1$ invests more than the mean-variance efficient allocation in assets that have a negative covariance with changes in the riskfree rate. These assets pay off when the riskfree rate is low, giving the investor more wealth when marginal utility for wealth is highest.

All of the reasoning above goes through regardless of the level of risk premia. Indeed, Figure 8 shows that hedging demand coming from the real riskfree rate does not depend on the value of the state variables. Mathematically, this follows from the fact that $B_3(\tau) \equiv 0$ when $\lambda_2^* = 0$, as noted in Section II. When the investor only hedges changes in the real interest rate, hedging demand is non-stochastic.

A number of studies (e.g., Brennan and Xia (2000), Sorensen (1999), Wachter (2003)) have argued that a time-varying riskfree rate leads investors with longer horizons to allocate a greater percentage of their portfolio to long-term bonds. According to this argument, long-term bonds are negatively correlated with the riskfree rate, and thus should be over-weighted in the portfolios of investors with risk aversion greater than one. The limitation of this argument is that it requires bonds to be real. Nominal bonds are negatively correlated with the *nominal* riskfree rate, but the investor desires to hedge the real riskfree rate $r - \pi$, and nominal bonds may not be negatively correlated with the real riskfree rate. For our calibration, long-term bonds are indeed negatively correlated with the real riskfree rate, though it is important to note that this is an empirical, not a theoretical result. Thus, the investor with risk aversion greater than one chooses to increase her allocation to long-term bonds relative to the myopic portfolio. Because changes in the real riskfree rate are persistent, the longer the investor's horizon, the greater the effect of the riskfree rate on indirect utility, and the greater is hedging demand.

We now consider the optimal allocation when the investor hedges bond risk premia, but not the riskfree rate (i.e., in the calculation of hedging demand, the real riskfree rate $r - \pi$ is assumed to be constant). This is calculated by setting $\zeta = \delta$ in the equations for B_3, B_2, and B_1. This allocation is shown in Figure 8 and marked with circles. When bond risk premia are positive, hedging demand induced by time-varying risk premia is also positive. When bond risk premia are negative, hedging demand induced by time-varying risk premia is negative at short horizons and positive at long horizons.

Consider first the case where risk premia on long-term bonds are positive. A rise in bond risk premia counts as an improvement in investment opportunities, while a fall in bond risk premia counts as a deterioration. The same reasoning that applied in the case of a time-varying riskfree rate applies here, too. When the income effect dominates ($\gamma > 1$), hedging demand is positive for assets that are negatively correlated with bond premia. As shown in Table V, bond prices and bond risk premia are negatively correlated. This explains why hedging demand for long-term bonds induced by time variation in risk premia is positive when bond risk premia are positive. Moreover, changes in risk premia are persistent, as are changes to the real riskfree rate. Thus, the longer is the investor's horizon, the greater is hedging demand, and the greater is the total allocation to the long-term bond.

This reasoning also explains why hedging demand for the long-term bond can be negative when bond risk premia are negative. Figure 8 shows that when risk premia are negative, hedging demand arising from time variation in risk premia causes the allocation to fall with the horizon before increasing again. When the investor is short the long-term bond, reductions in the risk premium represent improvements in the investment opportunity set. In order to hedge these changes, the investor has a more negative allocation to the long-term bond than the myopic investor. However, rather than steadily decreasing in the horizon, hedging demand begins to increase after a horizon of about 2 years, and eventually becomes positive.

This counter-intuitive result arises because average risk premia on bonds are positive. Because risk premia are linear functions of the mean-reverting state variables $X(t)$, risk premia that are negative in the present imply that, in the future, risk premia are likely to pass through zero. A long-horizon investor cares not only about risk premia today, but also about risk premia at all future points in time. Zero is the least advantageous value for the investor because neither a short nor a long position is profitable in expectation. All else equal, a long-term investor would prefer positive risk premia because they are more likely to stay positive, than negative risk premia are to stay negative. In a setting with time-varying equity premia, Campbell and Viceira (1999) and Kim and Omberg (1996) also note that when the risk premium is negative but close to zero, hedging demand is positive for long-horizon investors.

Returning to Figure 7, it is clear that optimal hedging demand is not simply a sum of hedging demand when only the real riskfree rate varies and hedging demand when only risk premia vary. It arises from a nonlinear interaction between the two. Because the investor uses the long-term bond to hedge time-variation in the real riskfree rate, she has an additional reason to prefer positive risk premia in the long run. This induces her to hedge risk premia to a greater extent than she would if the riskfree rate were constant.

This section has shown that accounting for time-variation in the risk premia on long-term bonds has two effects on the investor's optimal portfolio. First, it induces investors to time the bond market. A lower risk premium on a long-term bond leads the investor to allocate less wealth to the bond at all horizons. The second effect arises from the investor's wish to hedge changes in the risk premium and the real riskfree rate. This causes the optimal portfolio to increase dramatically with the horizon. Thus, the failure of the expectations hypothesis "matters" for long-term investors, at least in the case where the investor has access to a single long-term bond. The following section generalizes these results to the case where the investor has access to multiple long-term bonds.

B. Optimal Allocation to Multiple Long-Term Bonds

Panel A of Table VI shows the optimal allocation when the investor has access to a 3-year bond, a 10-year bond, stock, and a nominally riskless asset. As in the previous section, we determine the optimal allocation for the long-run mean of the state variables, and for the state variables plus and minus one standard deviation. We report only the effects of varying X_1.

For all three values of the state variable, the myopic portfolio consists of a short position in at least one of the bonds. These leveraged positions arise because of the correlation structure of bond returns implied by the model (and found in the data). Table IV shows the implied correlations in bond returns (Panel A), and correlations of monthly log bond returns from the data (Panel B).[26] As Table IV shows, bonds at all maturities are highly correlated. Thus, any estimated difference in the risk-return trade-off between the 3- and 10-year bonds leads the investor to leverage the bonds off one another. In the context of a time-varying real interest rate, Brennan and Xia (2002) and Campbell and Viceira (2001) also find that the investor takes highly levered positions in long-term bonds.

When risk premia are high and positive, the investor takes a leveraged position in the 10-year bond, financed by a short position in the 3-year bond and the riskfree asset. In this case, hedging demand makes the myopic allocation more extreme. Because the investor has a long position in the 10-year bond, decreases in the risk premium on the 10-year bond reflect deteriorations in the investment opportunity set. The investor hedges these changes in risk premia by allocating more, relative to myopic demand, to the 10-year bond. Because the investor has a short position in the 3-year bond, increases in the risk premium reflect deteriorations in the investment opportunity set. This leads the investor to take a greater short position in the 3-year bond. It is also the case

[26] Because yield data are unavailable for all maturities, the correlations in Panel B rely on approximating the yield on the 9-year, 11-month bond with the yield on the 10-year bond. Thus, the correlations in Panel B are essentially correlations between changes in yields.

Table VI

Optimal Allocation When There Are Multiple Long-Term Bonds

Panel A shows the optimal portfolios for different horizons when the investor is able to invest in the 3-year bond, 10-year bond, the stock, and the nominal riskfree asset. Panel B shows the optimal portfolios for different horizons when the investor is able to invest in the 1-year bond, 5-year bond, 10-year bond, the stock, and the nominal riskfree asset. Shown are the allocations to the bonds and the stock; the allocation to the riskfree asset is one minus the sum of these quantities. Allocations are shown for an investor with investment horizons of 0, 1, 10, and 20 years, and coefficients of relative risk aversion (γ) of 4 and 10. The state variables X_2 and X_3 are set equal to zero, while X_1 is set equal to minus one standard deviation (1.9) in the top panel, 0 in the middle panel, and plus one standard deviation in the bottom panel. "Premium" refers to the risk premium on the 5-year bond implied by each value of $X = (X_1, X_2, X_3)$.

		$\gamma = 4$				$\gamma = 10$			
		0 Years	1 Year	10 Years	20 Years	0 Years	1 Year	10 Years	20 Years
Panel A: Allocation to 2 Bonds and the Stock									
Premium = 12%	3-yr	−0.73	−1.64	−2.62	−2.67	−0.28	−0.70	−1.36	−1.55
	10-yr	2.79	3.39	3.85	3.93	1.11	1.41	1.80	1.99
	Stock	0.49	0.45	0.45	0.45	0.19	0.17	0.17	0.17
Premium = 2%	3-yr	2.86	4.13	4.66	4.61	1.16	1.86	2.11	1.93
	10-yr	−0.84	−1.11	−1.23	−1.15	−0.35	−0.48	−0.44	−0.26
	Stock	0.82	0.81	0.80	0.80	0.33	0.32	0.31	0.32
Premium = −8%	3-yr	6.46	9.90	11.94	11.89	2.59	4.41	5.58	5.40
	10-yr	−4.47	−5.62	−6.32	−6.23	−1.80	−2.37	−2.68	−2.51
	Stock	1.16	1.16	1.16	1.16	0.46	0.46	0.46	0.46
Panel B: Allocation to 3 Bonds and the Stock									
Premium = 12%	1-yr	28.61	19.53	19.30	19.42	11.44	7.22	7.46	7.54
	5-yr	−15.91	−13.28	−14.33	−14.48	−6.35	−5.21	−6.21	−6.52
	10-yr	8.09	7.83	8.55	8.70	3.22	3.16	3.84	4.15
	Stock	0.40	0.40	0.40	0.40	0.16	0.16	0.16	0.16
Premium = 2%	1-yr	21.58	19.22	19.50	19.63	8.62	7.51	7.62	7.69
	5-yr	−8.25	−6.66	−6.35	−6.51	−3.29	−2.51	−2.43	−2.72
	10-yr	2.29	1.75	1.61	1.77	0.90	0.65	0.73	1.02
	Stock	0.77	0.77	0.77	0.77	0.30	0.30	0.30	0.30
Premium = −8%	1-yr	14.54	18.91	19.71	19.85	5.81	7.80	7.77	7.84
	5-yr	−0.60	−0.03	1.63	1.46	−0.22	0.19	1.36	1.08
	10-yr	−3.51	−4.34	−5.33	−5.16	−1.42	−1.86	−2.39	−2.11
	Stock	1.13	1.13	1.13	1.13	0.45	0.45	0.45	0.45

that the 3-year bond acts as a hedge on the 10-year bond and vice versa; this may also contribute to the positions moving in opposite directions.

When risk premia are positive but closer to zero, the optimal allocation changes. Now the risk-return trade-offs are such that the myopic portfolio consists of a positive fraction of wealth in the 3-year bond and a negative fraction in the 10-year bond. Hedging demands also reverse in sign. For short horizons, hedging demand is positive for the 3-year bond and negative for the 10-year bond. At long horizons, however, hedging demand is positive for both the 10 and the 3-year bonds.

Finally, when risk premia are negative, the investor holds a positive position in the 3-year bond and a negative position in the 10-year bond. Hedging demands cause these positions to become more extreme. Investment opportunities deteriorate when the risk premium on the 10-year bond rises or the risk premium on the 3-year bond falls. The investor chooses the optimal portfolio so that wealth is higher when this occurs. That is, the investor increases her weight in the 3-year bond and decreases her weight in the 10-year bond.

Panel B of Table VI examines the case where the investor has access to three long-term bonds, stock, and the riskfree asset. Because the nominal market is complete in this last case, it does not matter for the investor's utility or wealth which three bonds are chosen. Thus, without loss of generality, we assume that the investor has access to 1-, 5-, and 10-year bonds. Moreover, the investment opportunity set can be fully hedged by trading in bonds, thus hedging demand for stock is identically zero. The caveat stated for the case where the investor has access to two bonds applies to an even greater extent in this case. Because the three bonds are so highly correlated, the investor can achieve (perceived) high Sharpe ratios while taking on less risk than when he had access to fewer bonds. This leads to a highly leveraged myopic portfolio.

When risk premia are positive, the myopic allocation consists of a positive position in the 10-year and 1-year bonds and a negative position in the 5-year bond. Hedging demand is non-monotonic for all three bonds. For the 1-year bond, hedging demand is negative, but is less negative for investors with longer horizons than for shorter horizons. For the 5-year bond, hedging demand is positive, but is also smaller in magnitude for longer horizons than shorter horizons. For $\gamma = 10$, hedging demand switches sign for the 5-year bond and becomes positive at long horizons. For the 10-year bond hedging demand is negative at very short horizons (1 year) but positive at long horizons.

When risk premia are negative, the myopic allocation is still positive for the 1-year bond, but the allocation to the 10-year bond is below that for the 5-year bond.[27] Hedging demand is positive for the 1- and 5-year bonds, and negative for the 10-year bond. Hedging demand is monotonic except at very long horizons (20 years) where it diminishes slightly. In this case, hedging demand results in a more extreme allocation to all three bonds.

C. Utility Costs of Suboptimal Strategies

In order to assess the economic importance of the failure of the expectations hypothesis, we calculate utility costs under strategies that fail to take it into account. Three suboptimal strategies are considered. For the first (and least) suboptimal strategy, the investor times the bond-market optimally (the mean-variance efficient portfolio varies over time) and optimally hedges the real interest rate. However, the investor does not hedge time-varying risk premia. As

[27] For the parameter values we consider, they are both negative. When risk premia become more negative, however, the allocation to the 10-year bond is negative and the allocation to the 5-year bond is positive.

discussed in Section A, the optimal portfolio rule when the investor follows this strategy takes the form

$$\hat{w}(t) = \alpha_0 + \alpha_1 X(t), \tag{48}$$

where

$$\alpha_0 = \frac{1}{\gamma}(\sigma\sigma^\top)^{-1}\sigma\lambda_1^* + \frac{\gamma-1}{\gamma}(\sigma\sigma^\top)^{-1}\big(\sigma\sigma_\Pi^\top\big) + \frac{1}{\gamma}(\sigma\sigma^\top)^{-1}\big(\sigma\sigma_X^\top\big)B_2^*(\tau)^\top$$
$$\alpha_1 = \frac{1}{\gamma}(\sigma\sigma^\top)^{-1}\sigma\lambda_2^*,$$

where $B_2^*(\tau)$ is given by (B8) in the case of complete nominal markets and (C2) in the case of incomplete nominal markets, with λ_2^* set equal to zero. Note that $B_3(\tau) = 0$ if $\lambda_2^* = 0$. For this strategy, myopic demand is a function of the state variables, but hedging demand is non-stochastic.

For the second strategy we consider, the investor fails to hedge both time-varying risk premia and the time-varying riskfree rate, but follows the optimal myopic strategy. This strategy also takes the form (48), where

$$\alpha_0 = \frac{1}{\gamma}(\sigma\sigma^\top)^{-1}\sigma\lambda_1^* + \frac{\gamma-1}{\gamma}(\sigma\sigma^\top)^{-1}\big(\sigma\sigma_\Pi^\top\big)$$
$$\alpha_1 = \frac{1}{\gamma}(\sigma\sigma^\top)^{-1}\sigma\lambda_2^*.$$

We refer to this as the *conditional myopic strategy*.

Finally, we consider a static mean-variance investor. This investor's allocation equals the unconditional mean-variance portfolio, namely (48), with

$$\alpha_0 = \frac{1}{\gamma}(\sigma\sigma^\top)^{-1}\sigma\lambda_1^* + \frac{\gamma-1}{\gamma}(\sigma\sigma^\top)^{-1}\big(\sigma\sigma_\Pi^\top\big)$$
$$\alpha_1 = 0_{3\times 3}.$$

This investor neither hedges time-variation in the investment opportunity set nor times the bond market. We refer to this as the *unconditional myopic strategy*.

To calculate utility costs, we solve for indirect utility (38) when the investor follows a strategy of the form (48). Because indirect utility is an expectation of future direct utility, it is a martingale and thus has zero drift. From the Markov property, it is a function of $W(t)$, $\Pi(t)$, $X(t)$, and the horizon. Thus, indirect utility corresponding to the strategy $\hat{w}(t)$ must satisfy the partial differential equation:

$$J_t + \mathcal{L}J = 0, \tag{49}$$

where $\mathcal{L}J$ is the infinitesimal generator of J given by

$$\begin{aligned}\mathcal{L}J &= J_W W(\hat{w}^\top(\mu - r\iota) + r) + J_X \mu_X + J_\Pi \Pi \pi \\ &\quad + J_{WX} W \sigma_X \sigma^\top \hat{w}(t) + J_{W\Pi} W \Pi \hat{w}^\top \sigma \sigma_\Pi^\top + J_{X\Pi} \Pi \sigma_X \sigma_\Pi^\top \\ &\quad + \frac{1}{2} J_{WW} W^2 \hat{w}^\top \sigma \sigma^\top \hat{w} + \frac{1}{2} J_{\Pi\Pi} \Pi^2 \sigma_\Pi \sigma_\Pi^\top + \frac{1}{2} \operatorname{tr}\left(J_{XX} \sigma_X \sigma_X^\top\right). \end{aligned} \tag{50}$$

For the cases where the allocation is linear in $X(t)$, the solution of (49) takes the same form as indirect utility when an investor follows an optimal strategy. Namely, when $\gamma > 1$:[28]

$$\hat{J}(W(t), \Pi(t), X(t), t, T) = \frac{1}{1-\gamma}\left(\frac{W(t)}{\Pi(t)}\right)^{1-\gamma} \hat{H}(X(t), t, T),$$

where $\hat{H}(X(t), t, T)$ is exponential quadratic. When $\gamma = 1$, indirect utility is quadratic. More details can be found in Appendix D.

We measure the utility costs of following a suboptimal strategy by calculating the percent of wealth the suboptimal investor would be willing to give away in return for being "allowed" to follow the optimal strategy. In other words, we solve for the quantity $\wp$ such that

$$\begin{aligned}&\frac{1}{1-\gamma}\left(\frac{W(0)}{\Pi(0)}\right)^{1-\gamma} \hat{H}(X(0), 0, T) \\ &\quad = \frac{1}{1-\gamma}\left(\frac{W(0)(1 - \wp(X(0), 0, T))}{\Pi(0)}\right)^{1-\gamma} H(X(0), 0, T). \end{aligned} \tag{51}$$

The left-hand side is the time-zero indirect utility of the investor who follows a suboptimal strategy and starts with wealth $W(0)$. The right-hand side is the time-zero indirect utility of the investor who follows an optimal strategy and starts with wealth $W(0)(1 - \wp)$. Note that while $\wp$ measures utility cost, the units are fractions of wealth, not utils.

It follows from (51) that

$$\wp(X(0), 0, T) = 1 - \left[\frac{H(X(0), 0, T)}{\hat{H}(X(0), 0, T)}\right]^{\frac{1}{\gamma-1}}.$$

When $\gamma > 1$, $H < \hat{H}$, implying that the investor would be willing to give up a positive percent of wealth in order to follow the optimal strategy. Because H and $\hat{H}$ are positive, $\wp$ lies between zero and one. The closer $\wp$ is to one, the more wealth the investor would be willing to give up to follow the optimal strategy, and thus the greater the cost to following the suboptimal strategy.

Figure 9 plots the percent of wealth ($100 \times \wp$) an investor following a suboptimal strategy would be willing to give up in order to follow an optimal strategy when a 5-year bond, the stock, and the nominal riskfree asset are available.

[28] The form of the value function for $\gamma = 1$ is discussed in Appendix D.

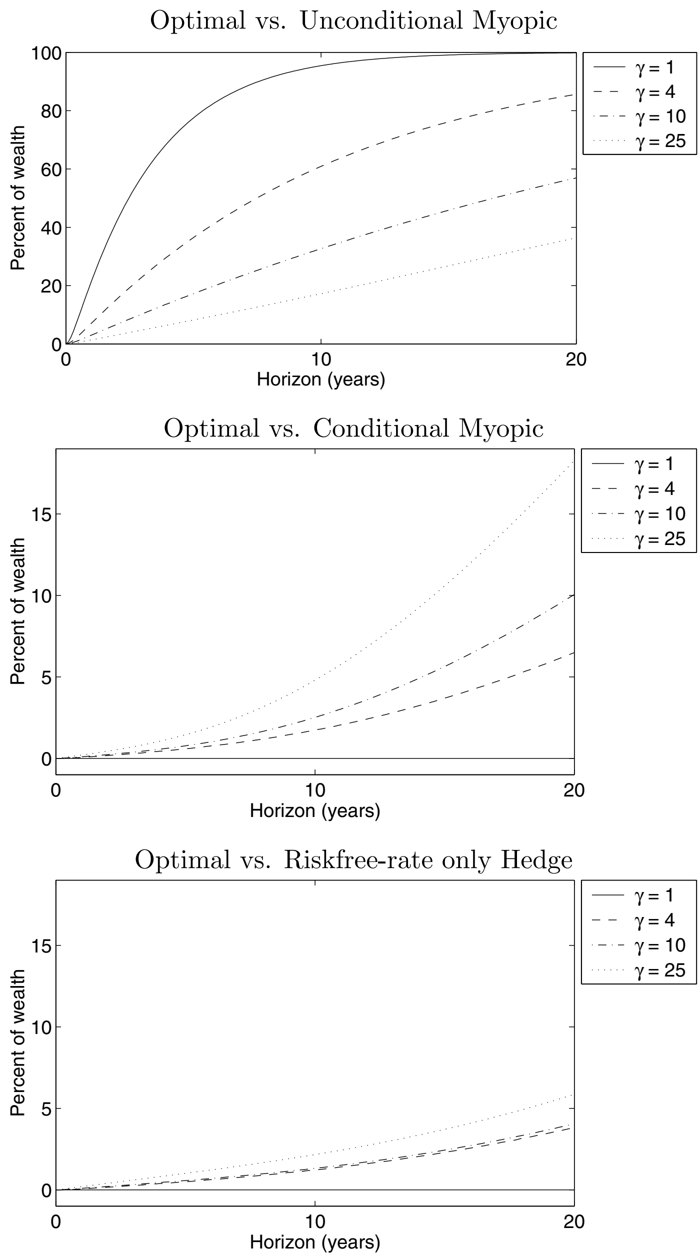

Figure 9. Utility costs as a function of horizon, when the investor has access to a 5-year bond, a stock, and a nominal riskfree asset. Utility costs are measured as the percent of wealth an investor following a suboptimal strategy would be willing to give up in order to follow an optimal strategy. The top panel shows the cost of following the unconditional myopic strategy. The middle panel shows the cost of following the conditional myopic strategy. The bottom panel shows the cost of following the conditional myopic strategy and hedging only the riskfree rate. The variable γ refers to relative risk aversion.

Utility costs are plotted for risk aversion equal to 1, 4, 10, and 25. The top panel plots the cost of following the unconditional myopic strategy, the middle panel the cost of following the conditional myopic strategy, and the lower panel the cost of hedging only the riskfree rate. For $\gamma > 1$, the costs fall as the strategies come closer to the optimum; the unconditional myopic strategy is more costly than the conditional myopic strategy, which is in turn more costly than only hedging the riskfree rate. For the log utility investor ($\gamma = 1$), it is optimal to do no hedging at all. Therefore, the costs of the conditional myopic strategy and the strategy of hedging only the riskfree rate are zero.

The cost of following the unconditional myopic strategy is very high. For $\gamma = 1$, an investor with a horizon of 20 years is willing to give up nearly 100% of wealth in order to follow the optimal strategy. The cost of this strategy falls as risk aversion rises. However, even the investor with risk aversion of 25 is willing to give up 40% of wealth in order to follow the optimal strategy. Failure to time and to hedge the bond market clearly results in a large utility loss for the investor.

Failure to hedge time-variation in investment opportunities is also costly. A highly risk averse investor ($\gamma = 25$) with a horizon of 20 years is willing to give up 20% of wealth in order to switch from the conditional myopic strategy to the optimal strategy. The cost of failing to hedge is lower, the lower is risk aversion. An investor with risk aversion equal to four is willing to give up 6% of wealth. Switching from a strategy that hedges only the riskfree rate to the fully optimal strategy is worth 6% of wealth for the investor with risk aversion of 25, and 4% of wealth for the investor with risk aversion of four. The cost of failing to hedge risk premia falls off more slowly with risk aversion than the overall cost of failing to hedge. This is because less risk averse investors allocate a greater percentage of their wealth to the long-term bond so they are more exposed to time-variation in risk premia.

Figure 10 shows the utility cost of suboptimal strategies when the investor has access to a 3-year and a 10-year bond, the stock, and the nominal riskfree asset. The utility costs of suboptimal strategies when two long-term bonds are available are higher than those when one long-term bond is available for all strategies and all investors. In particular, the costs of failing to hedge are much higher than before, and are no longer increasing in risk aversion. The investor with $\gamma = 4$ and a horizon of 20 years would give up 26% of wealth to switch from the conditional myopic strategy to the optimal strategy. The investor with $\gamma = 25$ would be willing to give up 24% of wealth to switch. Failure to hedge risk premia is nearly as costly as failure to hedge completely. The investor with $\gamma = 4$ is willing to give up 24% of wealth to switch from the strategy of hedging only the riskfree rate to the fully optimal strategy. In the case of two long-term bonds, there is greater scope for taking advantage of time-variation in risk premia. Because time-variation in risk premia is a more important component of hedging demand, and because less risk averse investors are particularly affected by time-variation in risk premia, the failure to hedge is costly even for investors with low risk aversion. It is also possible to compute utility costs of suboptimal strategies when the investor has access to three long-term bonds.

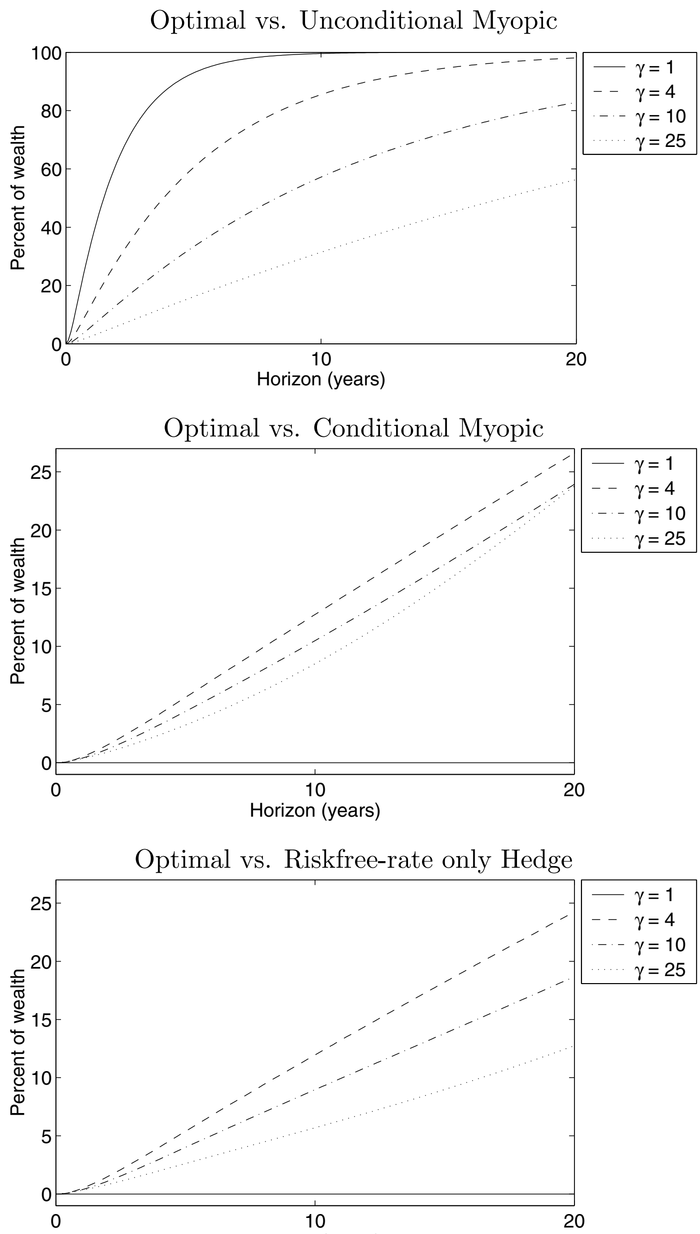

Figure 10. Utility costs as a function of horizon, when the investor has access to two long-term bonds, a stock, and a nominal riskfree asset. Utility costs are measured as the percent of wealth an investor following a suboptimal strategy would be willing to give up in order to follow an optimal strategy. The top panel shows the cost of following the unconditional myopic strategy. The middle panel shows the cost of following the conditional myopic strategy. The bottom panel shows the cost of following the conditional myopic strategy and hedging only the riskfree rate. The variable γ refers to relative risk aversion.

Results available from the authors show that in this case, the costs are larger still. Moreover, the cost of failing to hedge risk premia is even closer to the cost of failing to hedge at all.

This section has shown that following the unconditional myopic strategy, while optimal in the case where investment opportunities are constant, carries high utility costs when they are time-varying. The utility costs are large even when the investor allocates wealth between the nominally riskfree asset, one long-term bond, and the stock; the effect does not rely on the investor taking large offsetting positions in bonds of different maturities. Moreover, following a conditional myopic strategy is also costly, as is following a conditional myopic strategy but hedging only the riskfree rate (rather than both the riskfree rate and the risk premia). Thus, the failure of the expectations hypothesis is important for long-term investors: Treating risk premia as if they are constant results in economically significant costs.

V. Conclusion

We have shown that the failure of the expectations hypothesis has potentially important consequences for the portfolios of long-term investors. For an investor who allocates wealth between a long and a short-term bond, time-variation in risk premia induces hedging demand that is large and positive. We find that long horizon investors should hold a greater fraction of their portfolio in the long-term bond—an effect that persists out to a horizon of 20 years. When the investor has access to multiple long-term bonds, hedging demands generally make the optimal allocation more extreme. We find that failing to hedge time-variation in return predictability carries large utility costs for the long-term investor.

We establish these results by extending the affine term structure literature to account for expected inflation. Jointly estimating a process for inflation and bond prices produces a series for expected inflation that can account for a large portion of the variance of realized inflation, even though it is constructed from bond yields alone.

Our framework is rich enough to include time-variation in the real interest rate, in risk premia on stock returns, and in expected inflation, but at the same time admits explicit solutions in near-to-closed form. Multiple extensions of our model are possible. For example, it is possible to extend our empirical results to allow for state variables other than those extracted from bond yields. We could examine the relative importance of these state variables, as in Ait-Sahalia and Brandt (2001). We could also modify our model to allow for parameter uncertainty, as in Barberis (2000), or learning, as in Xia (2001). Clearly, there are important aspects of the portfolio choice problem that we do not address. Transaction costs, parameter uncertainty, and non-expected utility preferences have all been fruitfully explored in the context of stock-return predictability. Bonds present a similar, yet richer framework to explore these same issues.

Appendix A: Bond Prices

Following Cox, Ingersoll, and Ross (1985), we assume that bond prices are smooth functions of the state variables $X(t)$ and of time. No-arbitrage implies that P satisfies

$$P_X K(\theta - X(t)) + \frac{1}{2}\mathrm{tr}\big(P_{XX}\sigma_X\sigma_X^\top\big) + P_t - r(t)P = P_X\sigma_X\bar{\Lambda}(t) \tag{A1}$$

with boundary condition $P(X(t), t, t) = 1$. Equation (A1) follows from equating the instantaneous expected excess return to the volatility multiplied by the price of risk.

Conjecture that

$$P(X(t), t, T) = \exp\{A_2(\tau)X(t) + A_1(\tau)\}, \tag{A2}$$

where $\tau = T - t$. Substituting back into (A1) and matching coefficients on $X(t)$ and the constants produce the following system of ordinary differential equations for the row vector $A_2(\tau)$ and the scalar $A_1(\tau)$:

$$A_2'(\tau) = -A_2(\tau)(K + \sigma_X\bar{\lambda}_2) - \delta \tag{A3}$$

$$A_1'(\tau) = A_2(\tau)(K\theta - \sigma_X\bar{\lambda}_1) + \frac{1}{2}A_2(\tau)\sigma_X\sigma_X^\top A_2(\tau)^\top - \delta_0. \tag{A4}$$

The boundary conditions are $A_2(0) = 0_{1\times m}$ and $A_1(0) = 0$.

Appendix B: Optimal Portfolio Allocation

Proof of Theorem 1: It follows from the Markov property of (Π, Z_{ν^*}, X) that wealth under the optimal policy may be written as

$$W(t) = G(\Pi(t), Z_{\nu^*}(t), X(t), t, T)$$

for some function G. Define a function F such that

$$G(\Pi(t), Z_{\nu^*}(t), X(t), t, T) = \Pi(t)Z_{\nu^*}(t)^{\frac{1}{\gamma}}F(X(t), t, T).$$

Because wealth is an asset, it satisfies a no-arbitrage differential equation analogous to that of bonds. Applying Ito's lemma to G and matching the instantaneous expected excess return on wealth to its volatility times the price of risk produces:[29]

$$\mathcal{L}G + G_t - rG = \big(G_{Z_{\nu^*}}Z_{\nu^*}\big((\Lambda^* + \nu^*)^\top - \sigma_\Pi\big) + G_\Pi\Pi\sigma_\Pi + G_X\sigma_X\big)(\Lambda^* + \nu^*), \tag{B1}$$

[29] From Ito's lemma, we can write

$$dZ_{\nu^*}(t) = \mu_{Z_{\nu^*}}\,dt + \sigma_{Z_{\nu^*}}\,dz$$

with

$$\begin{aligned}\mu_{Z_{\nu^*}} &= \Big(r(t) - \pi(t) + (\Lambda^* + \nu^*)^\top(\Lambda^* + \nu^*) + \sigma_\Pi\sigma_\Pi^\top - \sigma_\Pi(\Lambda^* + \nu^*)\Big)Z_{\nu^*}(t)\\ \sigma_{Z_{\nu^*}} &= \Big((\Lambda^* + \nu^*)^\top - \sigma_\Pi\Big)Z_{\nu^*}(t).\end{aligned}$$

where

$$\mathcal{L}G = G_{Z_{\nu^*}}\mu_{Z_{\nu^*}} + G_\Pi \Pi\pi + G_X K(\theta - X) + G_{Z_{\nu^*}X}\sigma_X\sigma_{Z_{\nu^*}}^\top + \Pi G_{\Pi X}\sigma_X\sigma_\Pi^\top$$
$$+ \Pi G_{\Pi Z_{\nu^*}}\sigma_\Pi\sigma_{Z_{\nu^*}}^\top + \frac{1}{2}\big(G_{Z_{\nu^*}Z_{\nu^*}}\sigma_{Z_{\nu^*}}\sigma_{Z_{\nu^*}}^\top + \Pi^2 G_{\Pi\Pi}\sigma_\Pi\sigma_\Pi^\top + \mathrm{tr}\big(G_{XX}\sigma_X\sigma_X^\top\big)\big),$$

with boundary condition

$$G(\Pi(T), Z_{\nu^*}(T), X(T), T, T) = \Pi(T) Z_{\nu^*}(T)^{\frac{1}{\gamma}}.$$

Note that the no-arbitrage relationship for G only holds for the minimax pricing kernel ϕ_{ν^*}, while the bond pricing equation (A1) holds for any pricing kernel. Substituting (34) into (B1) results in the partial differential equation for F given in the text.

In order for optimal wealth to satisfy the dynamic budget constraint (25), the diffusion terms from the two processes must match. Therefore, the price of risk and the optimal portfolio must jointly satisfy:

$$\frac{1}{\gamma}(\Lambda^* + \nu^*)^\top + \frac{\gamma - 1}{\gamma}\sigma_\Pi + \frac{F_X}{F}\sigma_X = w^\top\sigma, \tag{B2}$$

where w is the $N \times 1$ vector of portfolio weights. The lefthand side follows from Ito's lemma applied to G. Inflation risk σ_Π is not spanned by the row vectors of σ; thus, for general ν, this equation does not have a solution.

We need to find ν^* so that the unhedgeable part of σ_Π drops out.[30] This is equivalent to setting the demand on the nontraded assets to zero. Rewrite σ_Π as

$$\sigma_\Pi = \big(\sigma_\Pi\sigma^\top\big)(\sigma\sigma^\top)^{-1}\sigma + \big(\sigma_\Pi - \big(\sigma_\Pi\sigma^\top\big)(\sigma\sigma^\top)^{-1}\sigma\big). \tag{B3}$$

The first term is the projection of σ_Π onto the traded assets. The second term is orthogonal to the traded assets. In order for (B2) to have a solution, ν^* must satisfy

$$\frac{1}{\gamma}(\nu^*)^\top = \frac{1-\gamma}{\gamma}\big(\sigma_\Pi - \big(\sigma_\Pi\sigma^\top\big)(\sigma\sigma^\top)^{-1}\sigma\big).$$

Therefore,

$$\nu^* = (1-\gamma)\big(\sigma_\Pi^\top - \sigma^\top(\sigma\sigma^\top)^{-1}\sigma\sigma_\Pi^\top\big)^\top. \tag{B4}$$

Because ν^* is orthogonal to the basis assets, $(\Lambda^* + \nu^*)$, where Λ^* is given by (15), is indeed a valid price of risk.

Substituting (B4) back into (B2) produces

$$\frac{1}{\gamma}(\mu - \iota r)^\top(\sigma\sigma^\top)^{-1}\sigma + \frac{\gamma - 1}{\gamma}\big(\sigma_\Pi\sigma^\top\big)(\sigma\sigma^\top)^{-1}\sigma + \frac{1}{F}F_X\big(\sigma_X\sigma^\top\big)(\sigma\sigma^\top)^{-1}\sigma = w^\top\sigma.$$

[30] The variable ν^* does not have to cancel out the unhedgeable parts of Λ^* because the columns of Λ^* are spanned by the rows of σ. In fact, this is the reason for defining Λ^* as a projection of $\bar{\Lambda}$ onto the available assets.

The equation for the optimal allocation (37) follows from multiplying both sides of the equation by $\sigma^\top(\sigma\sigma^\top)^{-1}$ and taking the transpose. Q.E.D.

Proof of Corollary 1: The argument follows that of Cox and Huang (1989), generalized to the case of unexpected inflation. The investor's problem at time 0 can equivalently be written as

$$\max_{W(t)>0} E_0[J(W(t), \Pi(t), X(t), t, T)],$$

subject to the static budget constraint. The first-order condition is given by

$$J_W(t) = l\,\phi_{\hat{\nu}}(t),$$

where $\phi_{\hat{\nu}}(t)$ is the minimax pricing kernel. We do not know a priori that $\phi_{\hat{\nu}} = \phi_{\nu^*}$. As is well known, the solution to (38) takes the form

$$J(W(t), \Pi(t), X(t), t, T) = \frac{1}{1-\gamma}\left(\frac{W(t)}{\Pi(t)}\right)^{1-\gamma} H(X(t), t, T). \tag{B5}$$

Our goal is to prove the relationship between the functions H and F.

Define $\hat{Z}$ analogously to (32) as

$$\hat{Z}(t) = (l\,\phi_{\hat{\nu}}(t)\Pi(t))^{-1}.$$

Then the investor's first-order condition can be re-written as

$$J_W(t) = \hat{Z}(t)^{-1}\Pi(t)^{-1}.$$

Substituting in from (B5) implies that

$$W(t) = \hat{Z}(t)^{\frac{1}{\gamma}}\Pi(t)H(X(t), t, T)^{\frac{1}{\gamma}}. \tag{B6}$$

Because $W(t)$ is an asset, it must satisfy partial differential equation (B1). Comparing (B6) with (34), it follows that $H^{\frac{1}{\gamma}}$ and $\hat{\nu}$ must jointly satisfy the partial differential equations (36) and (B2). Therefore, $\hat{\nu}$ must equal ν^*, and $H^{\frac{1}{\gamma}}$ must equal F. Q.E.D.

Proof of Corollary 3: To solve for F, we conjecture the form of it and then we verify. Our conjecture is that

$$F(X(t), t, T) = \exp\left\{\frac{1}{\gamma}\left(\frac{1}{2}X_t^\top B_3(\tau)X_t + B_2(\tau)X_t + B_1(\tau)\right)\right\},$$

where $\tau = T - t$, $B_3(\tau)$ is a matrix, $B_2(\tau)$ is a row vector, and $B_1(\tau)$ is a scalar. Substituting the hypothesized solution back into the PDE (36) and matching coefficients on $X(t)^\top[\,\cdot\,]X(t)$, $X(t)$, and the constant terms lead to a system of ordinary differential equations:

$$B_3'(\tau) = \left(B_3(\tau) + B_3(\tau)^\top\right)\left[\left(\frac{1}{\gamma} - 1\right)\sigma_X \lambda_2^* - K\right]$$
$$+ \left(\frac{1}{4\gamma}\left(B_3(\tau) + B_3(\tau)^\top\right)\sigma_X \sigma_X^\top \left(B_3(\tau) + B_3(\tau)^\top\right) + \left(\frac{1}{\gamma} - 1\right)\lambda_2^{*\top}\lambda_2^* \quad \text{(B7)}$$

$$B_2'(\tau) = B_2(\tau)\left[\left(\frac{1}{\gamma} - 1\right)\sigma_X \lambda_2^* - K + \frac{1}{2\gamma}\sigma_X \sigma_X^\top \left(B_3(\tau) + B_3(\tau)^\top\right)\right]$$
$$+ \frac{1}{2}\left[\theta^\top K^\top + \left(\frac{1}{\gamma} - 1\right)\lambda_1^{*\top}\sigma_X^\top + \left(1 - \frac{1}{\gamma}\right)\sigma_\Pi \sigma_X^\top\right]\left(B_3(\tau) + B_3(\tau)^\top\right)$$
$$+ (1-\gamma)(\delta - \zeta) + \left(\frac{1}{\gamma} - 1\right)\left(\lambda_1^* - \sigma_\Pi^\top\right)^\top \lambda_2^* - (\gamma - 1)\sigma_\Pi \lambda_2^* \quad \text{(B8)}$$

$$B_1'(\tau) = B_2(\tau)\left[K\theta + \left(\frac{1}{\gamma} - 1\right)\sigma_X \lambda_1^* + \left(1 - \frac{1}{\gamma}\right)\sigma_X \sigma_\Pi^\top\right]$$
$$+ \frac{1}{2\gamma}B_2(\tau)\sigma_X \sigma_X{}^\top B_2(\tau)^\top + \frac{1}{4}\mathrm{tr}\left(\left(B_3(\tau) + B_3(\tau)^\top\right)\sigma_X \sigma_X{}^\top\right)$$
$$+ \frac{1}{2}\left(\frac{1}{\gamma} - 1\right)\left[\left(\lambda_1^* + \nu^*\right)^\top - \sigma_\Pi\right]\left[\left(\lambda_1^* + \nu^*\right)^\top - \sigma_\Pi\right]^\top$$
$$+ (1-\gamma)\sigma_\Pi\left(\lambda_1^* + \nu^*\right) + (1-\gamma)(\delta_0 - \zeta_0). \quad \text{(B9)}$$

Q.E.D.

Appendix C: Optimal Portfolio Allocation under Incomplete Nominal Markets

This Appendix modifies the results above to the case where the investor has fewer bonds than state variables. In this case, nominal markets are incomplete. To determine the minimax price of risk in this case, we project σ_Π and σ_X on the available assets

$$\sigma_\Pi = \left(\sigma_\Pi \sigma^\top\right)(\sigma\sigma^\top)^{-1}\sigma + \left(\sigma_\Pi - \left(\sigma_\Pi \sigma^\top\right)(\sigma\sigma^\top)^{-1}\sigma\right)$$
$$\sigma_X = \left(\sigma_X \sigma^\top\right)(\sigma\sigma^\top)^{-1}\sigma + \left(\sigma_X - \left(\sigma_X \sigma^\top\right)(\sigma\sigma^\top)^{-1}\sigma\right).$$

It is useful to define the residual of the projections as

$$\left(\sigma_\Pi^\perp\right) = \sigma_\Pi - \left(\sigma_\Pi \sigma^\top\right)(\sigma\sigma^\top)^{-1}\sigma$$
$$\left(\sigma_X^\perp\right) = \sigma_X - \left(\sigma_X \sigma^\top\right)(\sigma\sigma^\top)^{-1}\sigma.$$

Equation (B2) and the same reasoning as in Appendix B implies that ν^* takes the form

$$\nu^* = (1-\gamma)\left(\sigma_\Pi - \left(\sigma_\Pi \sigma^\top\right)(\sigma\sigma^\top)^{-1}\sigma\right)^\top - \gamma\left(\sigma_X - \left(\sigma_X \sigma^\top\right)(\sigma\sigma^\top)^{-1}\sigma\right)^\top \frac{F_X^\top}{F}.$$

Conjecturing that the functional form for F is given by (43) in Corollary 3, we derive the following ODE's from (36):

$$\begin{aligned} B_3'(\tau) = \{\ \} &+ \frac{1}{4}\left(1 - \frac{1}{\gamma}\right)\left(B_3(\tau) + B_3(\tau)^\top\right)\left(\sigma_X^\perp\right)\left(\sigma_X^\perp\right)^\top\left(B_3(\tau) + B_3(\tau)^\top\right) \\ &- \left(\frac{1}{\gamma} - 1\right)\left(\frac{1}{2}\left(B_3(\tau) + B_3(\tau)^\top\right)\sigma_X\left(\sigma_X^\perp\right)^\top\left(B_3(\tau) + B_3(\tau)^\top\right)\right) \end{aligned} \tag{C1}$$

$$\begin{aligned} B_2'(\tau) = \{\ \} &+ \frac{1}{2}\left(1 - \frac{1}{\gamma}\right)\left(B_2\left(\sigma_X^\perp\right)\left(\sigma_X^\perp\right)^\top\left(B_3(\tau) + B_3(\tau)^\top\right)\right. \\ &\left.- B_2(\tau)\sigma_X\left(\sigma_X^\perp\right)^\top\left(B_3(\tau) + B_3(\tau)^\top\right)\right) \\ &+ \left(\frac{1}{\gamma} - 1\right)\left(\frac{1}{2}\left(B_2(\tau)\left(\sigma_X^\perp\right) - (1-\gamma)\left(\sigma_\Pi^\perp\right)\right)\left(\sigma_X^\perp\right)^\top\left(B_3(\tau) + B_3(\tau)^\top\right)\right) \\ &+ \frac{\gamma - 1}{2}\sigma_\Pi\left(\sigma_X^\perp\right)\left(B_3(\tau) + B_3(\tau)^\top\right) \end{aligned} \tag{C2}$$

$$\begin{aligned} B_1'(\tau) = \{\ \} &+ \frac{1}{2}\left(1 - \frac{1}{\gamma}\right)B_2\left(\sigma_X^\perp\right)\left(\sigma_X^\perp\right)^\top B_2^\top \\ &+ \frac{1}{2}\frac{1-\gamma}{\gamma}B_2(\tau)\left(\sigma_X^\perp\right)\left(\sigma_X^\perp\right)^\top B_2(\tau)^\top \\ &- \left(\frac{1}{\gamma} - 1\right)(1-\gamma)B_2(\tau)\left(\sigma_X^\perp\right)\left(\sigma_\Pi^\perp\right)^\top \\ &+ (\gamma - 1)\sigma_\Pi\left(\sigma_X^\perp\right)^\top B_2(\tau)^\top. \end{aligned} \tag{C3}$$

The terms { } represent the quantity on the right-hand side of equations (B7), (B8), and (B9), respectively.

Note that when markets are complete, the new terms on the right-hand side of (C1), (C2), and (C3) reduce to zero. In particular, $(\sigma_\Pi^\perp)\sigma_X^\top = 0$ because $\sigma_X^\top$ is now within the span of σ.

Appendix D: Indirect Utility for Suboptimal Strategies

It follows from the partial differential equation (49) that indirect utility takes the form

$$J(W(t), \Pi(t), X(t), T) = \frac{1}{1-\gamma}\left(\frac{W(t)}{\Pi(t)}\right)^{1-\gamma} H(X(t), t, T),$$

where $H(X(t), t, T)$ satisfies the partial differential equation

$$
\begin{aligned}
&H_t + (1-\gamma)H\Bigg(w(t)^\top(\mu(t) - \iota r(t)) + r(t) - \pi(t) \\
&\quad - \gamma\frac{1}{2}w(t)^\top\sigma\sigma^\top w(t) - (1-\gamma)\,w(t)^\top\sigma\sigma_\Pi^\top - \frac{\gamma-2}{2}\sigma_\Pi\sigma_\Pi^\top\Bigg) \\
&\quad + H_X\big(K(\theta - X(t)) + (1-\gamma)\sigma_X\sigma^\top w(t) - (1-\gamma)\sigma_X\sigma_\Pi^\top\big) + \frac{1}{2}\mathrm{tr}\big(H_{XX}\sigma_X\sigma_X^\top\big) = 0.
\end{aligned} \tag{D1}
$$

Suppose that the strategy of interest can be expressed as

$$
w(t) = \alpha_0 + \alpha_1 X(t), \tag{D2}
$$

for some constant scalar α_0 and vector α_1. When the trading strategy can be expressed as (D2), it follows from (D1) that $H(X(t), t, T)$ is exponential quadratic:

$$
H(X(t), t, T) = \exp\left\{\frac{1}{2}X(t)^\top\Gamma_3 X(t) + \Gamma_2 X(t) + \Gamma_1\right\},
$$

where Γ_3, Γ_2, and Γ_1 satisfy the following system of ordinary differential equations:

$$
\begin{aligned}
\Gamma_3' = {} & \big(\Gamma_3 + \Gamma_3^\top\big)\big[(1-\gamma)\sigma_X\sigma^\top\alpha_1 - K\big] + \frac{\Gamma_3 + \Gamma_3^\top}{2}\sigma_X\sigma_X^\top\frac{\Gamma_3 + \Gamma_3^\top}{2} \\
& + 2(1-\gamma)\alpha_1^\top\sigma\lambda_2^* - \gamma(1-\gamma)\alpha_1^\top\sigma\sigma^\top\alpha_1
\end{aligned} \tag{D3}
$$

$$
\begin{aligned}
\Gamma_2' = {} & \Gamma_2\Big[(1-\gamma)\sigma_X\sigma^\top\alpha_1 + \sigma_X\sigma_X{}^\top\frac{\Gamma_3 + \Gamma_3^\top}{2} - K\Big] \\
& + \big[\theta^\top K^\top + (1-\gamma)\alpha_0^\top\sigma\sigma_X^\top - (1-\gamma)\sigma_\Pi\sigma_X^\top\big]\frac{\Gamma_3 + \Gamma_3^\top}{2} \\
& + (1-\gamma)\big[\alpha_0^\top\sigma\lambda_2^* + \delta - \zeta + \big(\lambda_1^*\big)^\top\sigma^\top\alpha_1\big] - (1-\gamma)^2\sigma_\Pi\sigma^\top\alpha_1 - \gamma(1-\gamma)\alpha_0^\top\sigma\sigma^\top\alpha_1
\end{aligned} \tag{D4}
$$

$$
\begin{aligned}
\Gamma_1' = {} & \Gamma_2\big[K\theta + (1-\gamma)\sigma_X\sigma^\top\alpha_0 - (1-\gamma)\sigma_X\sigma_\Pi^\top\big] + \frac{1}{2}\Gamma_2\sigma_X\sigma_X{}^\top\Gamma_2^\top \\
& + (1-\gamma)\big(\alpha_0^\top\sigma\lambda_1^* + \delta_0 - \zeta_0\big) - (1-\gamma)^2\sigma_\Pi\sigma^\top\alpha_0 - \frac{\gamma(1-\gamma)}{2}\alpha_0^\top\sigma\sigma^\top\alpha_0 \\
& - \frac{(1-\gamma)(\gamma-2)}{2}\sigma_\Pi\sigma_\Pi^\top + \frac{1}{2}\mathrm{tr}\left(\frac{\Gamma_3 + \Gamma_3^\top}{2}\sigma_X\sigma_X^\top\right).
\end{aligned} \tag{D5}
$$

For $\gamma = 1$, the indirect utility function takes the form

$$
J(W(t), \Pi(t), X(t), T) = \log\left(\frac{W(t)}{\Pi(t)}\right) + Q(X(t), t, T),
$$

where $Q(X(t), t, T)$ satisfies the partial differential equation

$$Q_t + (w(t)^\top(\mu(t) - \iota r(t)) + r(t) - \pi(t)) - \frac{1}{2}w(t)^\top\sigma\sigma^\top w(t) + \frac{1}{2}\sigma_\Pi\sigma_\Pi^\top$$
$$+ Q_X(K(\theta - X(t))) + \frac{1}{2}\mathrm{tr}\big(Q_{XX}\sigma_X\sigma_X^\top\big) = 0. \tag{D6}$$

It follows from (D6) that $Q(X(t), t, T)$ is quadratic:

$$Q(X(t), t, T) = \frac{1}{2}X(t)^\top\Delta_3 X(t) + \Delta_2 X(t) + \Delta_1,$$

and that Δ_3, Δ_2, and Δ_1 satisfy:

$$\Delta_3' = -\big(\Delta_3 + \Delta_3^\top\big)K + 2\alpha_1^\top\sigma\lambda_2^* - \alpha_1^\top\sigma\sigma^\top\alpha_1$$
$$\Delta_2' = -\Delta_2 K + \theta^\top K^\top\frac{\Delta_3 + \Delta_3^\top}{2} + \alpha_0^\top\sigma\lambda_2^* + \delta - \zeta + \big(\lambda_1^*\big)^\top\sigma^\top\alpha_1 - \alpha_0^\top\sigma\sigma^\top\alpha_1$$
$$\Delta_1' = \Delta_2 K\theta + \left(\alpha_0^\top\sigma\lambda_1^* + \delta_0 - \zeta_0 - \frac{1}{2}\alpha_0^\top\sigma\sigma^\top\alpha_0\right) + \frac{1}{2}\sigma_\Pi\sigma_\Pi{}^\top$$
$$+ \frac{1}{2}\mathrm{tr}\left(\frac{\Delta_3 + \Delta_3^\top}{2}\sigma_X\sigma_X^\top\right).$$

Appendix E: Estimation

This section extends the results of Duffee (2002) to include inflation and stock return data in the estimation of bond yields. For convenience, it is assumed that the state variables are Gaussian (as in the body of the paper). Duffee's quasi-maximum likelihood results for square-root models can be extended in a similar fashion. In what follows, let e^M denote the matrix exponential of M, let $(x_i)_i$ denote the diagonal matrix with diagonal elements x_i, and let $(x_{i,j})_{i,j}$ denote the matrix with the (i,j) element equal to $x_{i,j}$. It is assumed that K is diagonalizable.

Let $Y(t)$ denote the vector of perfectly observed yields at time t. Namely

$$Y(t) = \begin{pmatrix} y(X(t), t, \tau_1) \\ \vdots \\ y(X(t), t, \tau_m) \end{pmatrix}$$

for maturities $(\tau_1, \ldots, \tau_m)$, where y is defined in (6). Let $\tilde{Y}$ denote the vector of yields which are observed imperfectly. From (5), it follows that the perfectly observed yields can be inverted to find the state variables

$$X(t) = L_1^{-1}(Y(t) - L_0),$$

where L_1 is an $m \times m$ matrix with row i given by $-A_2(\tau_i)/\tau_i$, and L_0 is a vector with elements $-A_1(\tau_i)/\tau_i$. Let $f(\cdot \mid \cdot)$ denote (with slight abuse of notation), the conditional likelihood function. Then the likelihood function for yields can be related to the likelihood function for the state variables by

$$
\begin{aligned}
&f(Y(t+1), S(t+1), \Pi(t+1) \mid Y(t), S(t), \Pi(t)) \\
&\quad = \frac{1}{\det[L_1]} f(X(t+1), S(t+1), \Pi(t+1) \mid X(t), S(t), \Pi(t)).
\end{aligned} \tag{E1}
$$

Let $\epsilon(t)$ denote the observation errors on the yields that are not perfectly observed. We assume that $\epsilon(t)$ is independent of innovations to the state variables or to inflation. Under this assumption, the full likelihood is given by

$$
\begin{aligned}
l_t &= \log f(Y(t), S(t), \Pi(t) \mid Y(t-1), S(t-1), \Pi(t-1)) \\
&\quad + \log f(\tilde{Y}(t) \mid Y(t), S(t), \Pi(t)) \\
&= \log f(Y(t), S(t), \Pi(t) \mid Y(t-1), S(t-1), \Pi(t-1)) + \log f(\epsilon(t) \mid Y(t)).
\end{aligned}
$$

It therefore suffices to specify $f(X(t+1), S(t+1), \Pi(t+1) \mid X(t), S(t), \Pi(t))$.

We show that $f(X(t+1), \log S(t+1), \log \Pi(t+1) \mid X(t), \log S(t), \log \Pi(t))$ is multivariate normal, and calculate the mean and variance. Consider the augmented state vector

$$
\hat{X}(t) = \begin{bmatrix} X(t) \\ \log S(t) \\ \log \Pi(t) \end{bmatrix}.
$$

The continuous time dynamics of this vector are defined by

$$
d\hat{X}(t) = (\kappa_1 \hat{X} + \kappa_2)\, dt + \sigma_{\hat{X}}\, dz, \tag{E2}
$$

where

$$
\kappa_1 = \begin{bmatrix} -K & 0_{3\times 1} & 0_{3\times 1} \\ \eta + \delta & 0 & 0 \\ \zeta & 0 & 0 \end{bmatrix}, \quad \kappa_2 = \begin{bmatrix} K\theta \\ \eta_0 + \delta_0 - \frac{1}{2}\sigma_S \sigma_S^\top \\ \zeta_0 - \frac{1}{2}\sigma_\Pi \sigma_\Pi^\top \end{bmatrix}, \quad \sigma_{\hat{X}} = \begin{bmatrix} \sigma_X \\ \sigma_S \\ \sigma_\Pi \end{bmatrix}.
$$

Applying Ito's lemma to the process $e^{-\kappa_1 t}\hat{X}_t$, it follows that

$$
\hat{X}(T) = e^{\kappa_1 (T-t)} \hat{X}_t + \int_t^T e^{\kappa_1 (T-s)} \kappa_2\, ds + \int_t^T e^{\kappa_1 (T-s)} \sigma_{\hat{X}}\, dw(s), \tag{E3}
$$

which shows that $\hat{X}_T$ is normally distributed conditional on $\hat{X}_t$.

Assume that κ_1 is diagonalizable. Let U be such that

$$
\kappa_1 = U D U^{-1}, \quad D \text{ diagonal.}
$$

From the definition of the matrix exponential and (E3), it follows that

$$
E_t(\hat{X}(T)) = e^{\kappa_1 (T-t)} \hat{X}(t) + \left(\int_t^T U e^{D(T-s)} U^{-1}\, ds \right) \kappa_2.
$$

Note that $e^{D(T-t)} = \left(e^{d_i(T-t)}\right)_i$. Performing the integration element-by-element produces

$$E_t(\hat{X}(T)) = e^{\kappa_1(T-t)}\hat{X}_t + U(f(d_i, T-t))_i U^{-1}\kappa_2,$$

where

$$f(d_i, T-t) = \begin{cases} -\dfrac{1}{d_i}\left(1 - e^{d_i(T-t)}\right) & d_i \neq 0 \\ T-t & d_i = 0. \end{cases}$$

This completes the derivation of the conditional mean.

From (E3), the conditional variance satisfies

$$\begin{aligned} \mathrm{Var}_t(\hat{X}(T)) &= E_t\left[\left(\int_t^T e^{\kappa_1(T-u)}\sigma_{\hat{X}}\, dw_u\right)\left(\int_t^T e^{\kappa_1(T-u)}\sigma_{\hat{X}}\, dw_u\right)^\top\right] \\ &= E_t\left[\int_t^T e^{\kappa_1(T-u)}\sigma_{\hat{X}}\sigma_{\hat{X}}^\top e^{\kappa_1(T-u)^\top} du\right] \\ &= \int_t^T e^{\kappa_1(T-u)}\sigma_{\hat{X}}\sigma_{\hat{X}}^\top \left(e^{\kappa_1(T-u)}\right)^\top du. \end{aligned}$$

Let $\Omega = U^{-1}\sigma_{\hat{X}}\sigma_{\hat{X}}^\top(U^{-1})^\top$. Integrating the above equation element-by-element produces

$$\begin{aligned} \mathrm{Var}_t(\hat{X}(T)) &= \int_t^T U e^{D(T-u)}\Omega e^{D(T-u)} U^\top\, du \\ &= U[g(d_i, d_j, T-t)\Omega_{i,j}]_{i,j} U^\top, \end{aligned}$$

where

$$g(d_i, d_j, T-t) = \begin{cases} -\frac{1}{d_i+d_j}\left(1 - e^{(d_i+d_j)(T-t)}\right) & d_i \neq 0 \text{ or } d_j \neq 0 \\ T-t & d_i = d_j = 0. \end{cases}$$

This completes the derivation of the conditional variance–covariance matrix.

REFERENCES

Ahn, Dong-Hyun, Robert F. Dittmar, and A. Ronald Gallant, 2002, Quadratic term structure models: Theory and evidence, *Review of Financial Studies* 15, 243–288.

Ait-Sahalia, Yacine, and Michael Brandt, 2001, Variable selection for portfolio choice, *Journal of Finance* 56, 1297–1351.

Ang, Andrew, and Geert Bekaert, 2003, The term structure of real rates and expected inflation, Working paper, Columbia University.

Backus, David, Silverio Foresi, Abon Mozumbar, and Liuren Wu, 2001, Predictable changes in yields and forward rates, *Journal of Financial Economics* 59, 281–311.

Balduzzi, Pierluigi, and Anthony W. Lynch, 1999, Transaction costs and predictability: Some utility cost calculations, *Journal of Financial Economics* 52, 47–78.

Bansal, Ravi, and Hao Zhou, 2002, Term structure of interest rates with regime shifts, *Journal of Finance* 57, 1997–2044.
Barberis, Nicholas, 2000, Investing for the long run when returns are predictable, *Journal of Finance* 55, 225–264.
Bekaert, Geert, and Robert J. Hodrick, 2001, Expectations hypothesis tests, *Journal of Finance* 56, 1357–1394.
Bekaert, Geert, Robert J. Hodrick, and David A. Marshall, 1997, On biases in tests of the expectations hypothesis of the term structure of interest rates, *Journal of Financial Economics* 44, 309–348.
Bekaert, Geert, Robert J. Hodrick, and David A. Marshall, 2001, Peso problem explanations for term structure anomolies, *Journal of Monetary Economics* 48, 241–270.
Boudoukh, Jacob, 1993, An equilbrium model of nominal bond prices with inflation-output correlation and stochastic volatility, *Journal of Money, Credit, and Banking* 25, 636–665.
Boudoukh, Jacob, Matthew Richardson, and Robert F. Whitelaw, 1994, Industry returns and the Fisher effect, *Journal of Finance* 49, 1595–1615.
Brandt, Michael W., 1999, Estimating portfolio and consumption choice: A conditional Euler equations approach, *Journal of Finance* 54, 1609–1645.
Brandt, Michael W., and David A. Chapman, 2002, Comparing multifactor models of the term structure, Manuscript, University of Texas and the Wharton School.
Brennan, Michael J., Eduardo S. Schwartz, and Ronald Lagnado, 1997, Strategic asset allocation, *Journal of Economic Dynamics and Control* 21, 1377–1403.
Brennan, Michael J., and Yihong Xia, 2000, Stochastic interest rates and the bond-stock mix, *European Finance Review* 4, 197–210.
Brennan, Michael J., and Yihong Xia, 2002, Dynamic asset allocation under inflation, *Journal of Finance* 57, 1201–1238.
Campbell, John Y., 1986, A defense of traditional hypotheses about the term structure of interest rates, *Journal of Finance* 41, 183–193.
Campbell, John Y., Y. Lewis Chan, and Luis M. Viceira, 2003, A multivariate model of strategic asset allocation, *Journal of Financial Economics* 67, 41–80.
Campbell, John Y., and Robert J. Shiller, 1988, The dividend-price ratio and expectations of future dividends and discount factors, *Review of Financial Studies* 58, 495–514.
Campbell, John Y., and Robert J. Shiller, 1991, Yield spreads and interest rate movements: A bird's eye view, *Review of Economic Studies* 58, 495–514.
Campbell, John Y., and Luis M. Viceira, 1999, Consumption and portfolio decisions when expected returns are time-varying, *Quarterly Journal of Economics* 114, 433–495.
Campbell, John Y., and Luis Viceira, 2001, Who should buy long-term bonds? *American Economic Review* 91, 99–127.
Cochrane, John, and Monika Piazzesi, 2002, Bond risk premia, Working paper #9178, NBER.
Cox, John C., and Chi-Fu Huang, 1989, Optimal consumption and portfolio policies when asset prices follow a diffusion process, *Journal of Economic Theory* 39, 33–83.
Cox, John C., Jonathan C. Ingersoll, and Stephen A. Ross, 1981, A re-examination of traditional hypotheses about the term tructure of interest rates, *Journal of Finance* 36, 769–799.
Cox, John C., Jonathan C. Ingersoll, and Stephen A. Ross, 1985, A theory of the term structure of interest rates, *Econometrica* 53, 385–408.
Dai, Qiang, and Kenneth Singleton, 2000, Specification analysis of affine term structure models, *Journal of Finance* 55, 1943–1978.
Dai, Qiang, and Kenneth Singleton, 2002, Expectations puzzles, time-varying risk premia, and affine models of the term structure, *Journal of Financial Economics* 63, 415–442.
Dai, Qiang, and Kenneth Singleton, 2003, Term structure dynamics in theory and reality, *Review of Financial Studies* 16, 631–678.
Duffee, Gregory R., 2002, Term premia and interest rate forecasts in affine models, *Journal of Finance* 57, 405–443.
Duffie, Darrell, and Rui Kan, 1996, A yield-factor model of interest rates, *Mathematical Finance* 6, 379–406.

Dybvig, Philip H., and Chi-Fu Huang, 1988, Nonnegative wealth, absence of arbitrage, and feasible consumption plans, *Review of Financial Studies* 1, 377–401.

Fama, Eugene F., 1975, Short-term interest rates as predictors of inflation, *American Economic Review* 65, 269–282.

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, 1989, Business conditions and expected returns on stocks and bonds, *Journal of Financial Economics* 29, 23–49.

Fama, Eugene F., and Michael R. Gibbons, 1982, Inflation, real returns, and capital investment, *Journal of Monetary Economics* 13, 297–323.

Fisher, Mark, 1998, A simple model of the failure of the expectations hypothesis, Working paper, Federal Reserve Bank of Atlanta.

Gibbons, Michael R., and Krishna Ramaswamy, 1993, A test of the Cox, Ingersoll, Ross model of the term structure, *Review of Financial Studies* 6, 619–658.

He, Hua, and Neil D. Pearson, 1991, Consumption and portfolio policies with incomplete markets and short-sale constraints: The infinite dimensional case, *Journal of Economic Theory* 54, 259–304.

Karatzas, Ioannis, John P. Lehoczky, and Steven E. Shreve, 1987, Optimal portfolio and consumption decisions for a small investor on a finite horizon, *SIAM Journal of Control and Optimization* 25, 1557–1586.

Kim, Tong Suk, and Edward Omberg, 1996, Dynamic nonmyopic portfolio behavior, *Review of Financial Studies* 9, 141–161.

Liu, Jun, 1999, Portfolio selection in stochastic environments, Working paper, Stanford University.

Liu, Jun, and Jun Pan, 2003, Dynamic derivative strategies, *Journal of Financial Economics* 69, 401–430.

Longstaff, Francis A., 2000a, Arbitrage and the expectations hypothesis, *Journal of Finance* 55, 989–994.

Longstaff, Francis A., 2000b, The term structure of very short-term rates: New evidence for the expectations hypothesis, *Journal of Financial Economics* 58, 397–415.

Merton, Robert C., 1971, Optimum consumption and portfolio rules in a continuous-time model, *Journal of Economic Theory* 3, 373–413.

Mishkin, Frederic S., 1981, The real rate of interest: An empirical investigation, *Carnegie-Rochester Series on Public Policy* 15, 1951–200.

Pennacchi, George G., 1991, Identifying the dynamics of real interest rates and inflation: Evidence using survey data, *Review of Financial Studies* 4, 53–86.

Piazzesi, Monika, 2004, Affine term structure models, in Yacine Ait-Sahalia and Lars Peter Hansen, eds.: *Handbook of Financial Econometrics*. (forthcoming).

Pliska, Stanley R., 1986, A Stochastic calculus model of continuous trading: Optimal portfolios, *Mathematics of Operations Research* 11, 239–246.

Schroder, Mark, and Costis Skiadas, 1999, Optimal consumption and Portfolio Selection with Stochastic Differential Utility, *Journal of Economic Theory* 89, 68–126.

Schroder, Mark, and Costis Skiadas, 2002, Optimal lifetime consumption-portfolio strategies under trading constraints and generalized recursive preferences, Working paper #285, Northwestern University.

Sorensen, Carsten, 1999, Dynamic asset allocation and fixed income management, *Journal of Financial and Quantitative Analysis* 34, 513–532.

Stambaugh, Robert, 1999, Predictive regressions, *Journal of Financial Economics* 54, 375–421.

Sun, Tong-Sheng, 1992, Real and nominal interest rates: A discrete-time model and its continuous-time limit, *Review of Financial Studies* 5, 581–611.

Valkanov, Rossen, 1998, The term structure with highly persistent interest rates, Working paper, University of California, Los Angeles.

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

Wachter, Jessica A., 2002, Portfolio and consumption decisions under mean-reverting returns: An exact solution for complete markets, *Journal of Financial and Quantitative Analysis* 37, 63–91.

Wachter, Jessica A., 2003, Risk aversion and allocation to long-term bonds, *Journal of Economic Theory* 112, 325–333.

Xia, Yihong, 2001, Learning about predictability: The effects of parameter uncertainty on dynamic asset allocation, *Journal of Finance* 56, 205–246.

Xia, Yihong, 2002, Long-term bond markets and investor welfare, Working paper, University of Pennsylvania.