Stock and Bond Pricing in an Affine Economy

Geert Bekaert and Steven R. Grenadier*

October 2001

*Bekaert is at the Columbia Business School, Columbia University, and Grenadier is at the Graduate School of Business, Stanford University. We thank Andrew Ang, Hendrik Bessembinder, Yael Hochberg, Bob Hodrick, Martin Lettau, Jun Liu and Jay Shanken for their helpful comments. We also thank seminar participants at the AFA meetings in Boston, Arizona State University, Erasmus Universiteit Rotterdam, the Federal Reserve Bank of New York, Harvard Business School, London Business School, New York University, Rijksuniversiteit Limburg in Maastricht, Rochester, Stanford University, Tilburg University, UCLA, Universiteit Gent, University of Iowa, University of North Carolina, Universiteit van Amsterdam and Vanderbilt for helpful comments, and Financial Engines, Inc. for their support on this paper. Geert Bekaert thanks the NSF for financial support. Correspondence can be addressed to Geert Bekaert, Columbia Business School, Uris Hall, Room 802, New York, New York 10027. Phone: (212) 854-9156, E-mail: gb241@columbia.edu.

Stock and Bond Pricing in an A¢ne Economy

October 2001

Abstract

This article provides a stochastic valuation framework for bond and stock returns that builds on three di¤erent pricing traditions: a¢ne models of the term structure, present-value pricing of equities, and consumption-based asset pricing. Our model provides a more general application of the a¢ne framework in that both bonds and equities are priced in a consistent fashion. This pricing consistency implies that term structure variables help price stocks while stock price fundamentals help price the term structure. We illustrate our model by considering three examples that are similar in spirit to well-known pricing models that fall within our general framework: a Mehra and Prescott (1985) economy, a present value model similar to Campbell and Shiller (1988), and a model with stochastic risk aversion similar to Campbell and Cochrane (1999). The empirical performance of our models is explored, with a particular emphasis on return predictability.

1 Introduction

The pricing of bonds and equities has mostly evolved along separate lines in the finance literature. Models of the term structure of interest rates are typically silent on the pricing of equities, while models of equity pricing are typically silent on the pricing of bonds. This state of affairs is quite surprising from both a theoretical and empirical vantage point. Given that there is no obvious reason for market segmentation between bond and equity markets, an internally consistent model should use one pricing model for both asset classes. Variables that determine the pricing of equity should also determine the pricing of bonds, and vice versa. For example, dividend growth rates should help determine the term structure, while the term spread should help determine the equity return.

The literature on the term structure of interest rates, evolving out of Cox, Ingersoll and Ross (1985), has been widely applied to the pricing of bonds and interest rate dependent derivative securities. Much of the most widely-known term structure models fall into what is known as the affine class, a particularly tractable class of models in which the yield on any zero coupon bond can be written as an affine function of the set of state variables.[1] Equity pricing models have typically fallen into two classes: present value pricing models and equilibrium pricing models. Present value models, such as the famous Gordon model, discount an infinite stream of dividends using an exogenous discount rate. Dynamic versions of such models allow for complex dividend processes and a time-varying discount rate.[2] Equilibrium equity pricing models, such as Lucas (1978) and Mehra and Prescott (1985), allow for utility maximization, production opportunities, and market clearing to result in equilibrium equity prices.

Empirically, there is substantial evidence that matching equity and bond return moments simultaneously might provide very powerful tests. From the equity premium literature [see especially Weil (1989)], it is apparent that the equity premium puzzle is as much a low risk-free rate puzzle as it is a high expected stock return puzzle. From the work of Fama and French (1989) and Keim and Stambaugh (1986), we know that there are common predictable components in bond and equity returns.

This paper makes several contributions to the asset pricing literature. First, we formulate a general pricing model, consisting of a pricing kernel and a set of state variables. The specified dynamics imply a set of arbitrage-free asset prices for bonds and equities. The term structure is affine, whereas equity price-dividend ratios are equal to the sum of exponentials of an affine function of the state variables. Importantly, a subset of the set of state variables represents observable economic factors

[1] Examples include the models of Vasicek (1977), Cox, Ingersoll, and Ross (1985), Ho and Lee (1986), and Pearson and Sun (1994). Duffie and Kan (1996) provide necessary and sufficient conditions under which an affine term structure model is consistent with the absence of arbitrage. Dai and Singleton (2000) provides a detailed empirical analysis of affine term structure models.

[2] See Campbell and Shiller (1988), Bollerslev and Hodrick (1996), Cochrane (1992), and more recently Ang and Liu (1999), Bakshi and Chen (1998) and Berk, Green and Naik (1999).

such as dividend growth and inflation. This will be critical to both the interpretation and empirical identification of the model. The remaining state variables represent unobserved (or difficult to measure) factors such as productivity shocks or stochastic risk aversion.[3] The pricing kernel dynamics encompass a very wide class of models, both models with exogenous discount rates, and models where the pricing kernel is fully determined by the first order conditions of an equilibrium model, and hence represents an intertemporal marginal rate of substitution. A similar approach is found in the SAINTS model of Constantinides (1992), which provides a continuous-time pricing kernel framework that can be supported in equilibrium. However, the SAINTS model is used only to price bonds, and the equilibrium links are not explored.

Our second contribution is that our general model makes precise the nature of equilibrium restrictions of a large class of models. Specifically, the pricing model embeds equilibrium models of equity pricing as well as dividend discount models with exogenous discount rates. These features are made clear as we work through specific examples of the general model using a three factor model. For example, when we illustrate our model with a prototypical example of a dividend discount model with a time-varying exogenous discount rate [e.g., Campbell and Shiller (1988)], we demonstrate that a special case of the model is an equilibrium economy in the Lucas (1978) and Mehra and Prescott (1985) tradition. We also demonstrate how our framework embeds a consumption-based asset pricing model similar to Campbell and Cochrane (1999), which permits cyclical variation in risk aversion, and thus in the risk premia on risky assets. Such models provide a potential explanation for the observed empirical phenomenon that equity risk premia are larger during economic downturns than during economic expansions.

Our third contribution is to examine the empirical performance of three versions of the model. We estimate the structural model parameters using GMM [Hansen (1982)], in such a way that each model matches salient features of the fundamental processes (dividend growth rates and inflation), the short rate process, and the equity premium. Note that, in order to compare our model with present value models, we do not use consumption growth rates to calibrate it to the data. We then over-identify the model looking at other moments for equity and bond returns and their covariance. However, our main focus is on return predictability. We investigate endogenous predictability by computing variance ratios, regression coefficients of returns on instruments such as dividend yields and term spreads, and by characterizing the conditional risk premiums implied by the models. Finally, we revisit the excess volatility puzzle by computing the variability of price-dividend ratios implied by the various models.

Our paper is organized as follows. Section 2 presents the general affine model structure and the pricing of bonds and equities. In Section 3, we specialize to a three-factor variant of the model and provide specific examples that fall within the affine framework. The models provided encompass both equilibrium and exogenous

[3] Ang and Piazzesi (2000) also explore the importance of economic factors in affine term structure models.

discounting models, and the connections between the two modeling approaches are made clear. Section 4 discusses the estimation strategy for the three-factor models, presents parameter estimates and compares some unconditional moments implied by the models with the data. Section 5 examines the endogenous predictability of returns. Given the increasingly worrisome evidence on statistical biases in small samples, we are careful to distinguish small sample from population behavior. Section 6 concludes.

2 The General Model

In this section we specify the dynamics of the underlying sources of uncertainty in the economy and of the pricing kernel process. We then use these specifications to derive the pricing equations for bonds and equities. The resulting pricing equations will fall within an affine class of models. That is, the term structure of interest rates will be equal to an affine function of the underlying state variables. Similarly, the pricing structure of equities will fall within what one might refer to as an "exponential-affine" class. Specifically, the price-dividend ratio will equal a sum of elements, where the log of each element is an affine function of the underlying state variables.[4]

Consider an economy with N state variables that summarize the fundamental uncertainty of the economy. Let Y_t be the N-dimensional vector of state variables, with $Y_t^{'} = (Y_{1,t}, Y_{2,t}, ..., Y_{N,t})$. A subset of the N state variables represents observable economic factors such as dividend growth and inflation, while the remaining state variables represent unobserved (or difficult to measure) factors such as productivity shocks or stochastic risk aversion. One of the elements of the vector Y_t will always represent real dividend growth, Δd_t, and one element will always represent inflation, π_t. Thus, if D_t represents the real level of aggregate dividends and Λ_t represents the price level, then $\Delta d_{t+1} = \ln(D_{t+1}/D_t)$ and $\pi_{t+1} = \ln(\Lambda_{t+1}/\Lambda_t)$. The additional state variables will vary under different specifications. Let $\shortparallel \cdot \shortparallel$ denote the function defined by:

$$\shortparallel Y_{j,t} \shortparallel \equiv \begin{cases} \sqrt{Y_{j,t}} & \text{if } Y_{j,t} \geq 0 \\ 0 & \text{if } Y_{j,t} < 0. \end{cases} \tag{1}$$

Let F_t denote the $N \times N$ diagonal matrix with the elements $(\shortparallel Y_{1,t} \shortparallel, \shortparallel Y_{2,t} \shortparallel, ..., \shortparallel Y_{N,t} \shortparallel)$ along the diagonal. Writing this in matrix form,

$$F_t = (\shortparallel Y_{1,t} \shortparallel, \shortparallel Y_{2,t} \shortparallel, ..., \shortparallel Y_{N,t} \shortparallel)^{'} \odot I, \tag{2}$$

where I is the identity matrix of order N, and $\odot$ denotes the Hadamard Product.[5]

[4] Our framework does not (nor is it intended to) represent the most general structure for providing affine bond yields and exponential-affine price-dividend ratios. Our framework is, however, sufficiently general to embed many well-known asset pricing models, as illustrated in this paper. For a more general affine framework, one should refer to Duffie and Kan (1996).

[5] The Hadamard Product operator denotes element-by-element multiplication. We define it for-

The dynamics of Y_t follow a simple, first-order vector autoregressive (VAR) stochastic process:

$$Y_{t+1} = \mu + AY_t + (\Sigma_F F_t + \Sigma_H)\,\varepsilon_{t+1}, \tag{3}$$

with $\varepsilon_{t+1} \sim N(0, I)$ representing the fundamental shocks to the economy. The time t conditional expected value of Y_{t+1} is equal to $\mu + AY_t$, where μ is an N-dimensional column vector and A is an $N \times N$ matrix. The time t conditional volatility of Y_{t+1} is represented by $\Sigma_F F_t + \Sigma_H$, where Σ_F and Σ_H are $N \times N$ matrices representing sensitivities to the fundamental economic shocks.

In essence, the dynamics of Y_t represent a discrete-time system of a multidimensional combination of Vasicek and square-root processes. For example, if A and Σ_F are diagonal, and $\Sigma_H = 0$, Y_t would contain N square-root processes. Similarly, if A and Σ_H are diagonal, and $\Sigma_F = 0$, Y_t would contain N AR(1) processes.

Given the specification of the dynamics of Y_t, the pricing model is completed by specifying a pricing kernel (or stochastic discount factor). The (real) pricing kernel, M_t, is a positive stochastic process that ensures that all assets i are priced such that:

$$1 = E_t\left[(1 + R_{i,t+1})\, M_{t+1}\right], \tag{4}$$

where $R_{i,t+1}$ is the percentage real return on asset i over the period from t to $t+1$, and E_t denotes the expectation conditional on the information at time t. The existence of such a pricing kernel is ensured in any arbitrage-free economy. Harrison and Kreps (1979) derive the conditions under which M_t is unique. Let $m_{t+1} = \ln(M_{t+1})$.

The log of the real pricing kernel is specified as:

$$m_{t+1} = \mu_m + \Gamma_m^{'} Y_t + \left(\Sigma_{mf}^{'} F_t + \Sigma_m^{'}\right) \varepsilon_{t+1}, \tag{5}$$

where Γ_m, Σ_{mf}, and Σ_m are N-dimensional column vectors, and μ_m is a scalar.

In order to price nominally denominated assets, we must work with a nominal pricing kernel. Let the nominal pricing kernel be denoted by $\hat{m}_{t+1}$. The nominal pricing kernel is simply the real pricing kernel minus inflation: $\hat{m}_{t+1} = m_{t+1} - \pi_{t+1}$.[6]

mally in the Appendix. A useful implication of the Hadamard Product is that if $Y_{j,t} \geq 0, \forall j$, then $F_t F_t^{'} = Y_t \odot I$.

[6] This is simple to demonstrate. Let P_t denote the real price of an asset at time t, and let D_{t+1} denote its real payout at time $t+1$. Let Λ_t denote the price level. The nominal price of the asset is simply $P_t \Lambda_t \equiv P_t^N$ and its nominal payout is $D_{t+1}\Lambda_{t+1} \equiv D_{t+1}^N$. Using the real kernel, the real price may be expressed as:

$$P_t = E_t\left[(P_{t+1} + D_{t+1}) M_{t+1}\right].$$

Rewriting the above expression:

$$\begin{aligned} P_t \Lambda_t &= E_t\left[(P_{t+1}\Lambda_{t+1} + D_{t+1}\Lambda_{t+1})\left(\frac{\Lambda_t}{\Lambda_{t+1}}\right) M_{t+1}\right] \\ P_t^N &= E_t\left[(P_{t+1}^N + D_{t+1}^N)\left(\frac{\Lambda_t}{\Lambda_{t+1}}\right) M_{t+1}\right] \end{aligned}$$

In order to ensure that the specification of the process Y_{t+1} and m_{t+1} permits a well-defined system of pricing equations, as well as ensuring that the resulting pricing system falls within the affine class, we impose the following four restrictions on the processes:

$$\begin{aligned} \Sigma_F F_t \Sigma_H^{'} &= \mathbf{0}, \\ \Sigma_{mf}^{'} F_t \Sigma_m &= 0, \\ \Sigma_H F_t \Sigma_{mf} &= \mathbf{0}, \\ \Sigma_F F_t \Sigma_m &= \mathbf{0}. \end{aligned} \tag{6}$$

The main purpose of these restrictions is to exclude certain mixtures of square-root and Vasicek processes in the state variables and pricing kernel that lead to an intractable solution.

We can now combine the specification for Y_t and m_{t+1} to price financial assets. The details of the derivations are presented in the Appendix. It is important to note that, due to the discrete-time nature of the model, these solutions only represent approximate solutions to the true asset prices. The nature of the approximation results from the fact that if one of the state variables can become negative, and if the specific model allows for a stochastic volatility term containing a square-root process, we must rely on the $\| \cdot \|$ function to make the square-root well defined. When the state variable is then forced to reflect at zero, our use of the conditional lognormality features of the state variables becomes incorrect. However, this effect is minimized in the following ways. First, the square-root process is not always utilized in some of the standard applications of our model, in which case the pricing formulas are exact. Second, even in the case in which a state variable is forced to reflect at zero, reasonable parameterizations of the model can ensure that the likelihood of such a reflection is quite small. Finally, the exact solution can be computed numerically (for example, using quadrature), which would overcome the analytical approximation, but would also introduce approximation error. For these reasons, we have decided to present the simple affine solutions both to ensure the tractability of the results, and because of the close approximation in most instances.

Let us begin by deriving the pricing of the nominal term structure of interest rates. Let the time t price for a default-free zero-coupon bond with maturity n be denoted by $P_{n,t}$. Using the nominal pricing kernel, the value of $P_{n,t}$ must satisfy:

$$P_{n,t} = E_t \left[\exp\left(\hat{m}_{t+1}\right) P_{n-1,t+1} \right], \tag{7}$$

where $\hat{m}_{t+1} = m_{t+1} - \pi_{t+1}$ is the log of the nominal pricing kernel. Let $p_{n,t} = \ln(P_{n,t})$. The n-period bond yield is denoted by $y_{n,t}$, where $y_{n,t} = -p_{n,t}/n$. The solution to

Thus,

$$1 = E_t \left[\left(\frac{P_{t+1}^N + D_{t+1}^N}{P_t^N} \right) \exp(m_{t+1} - \pi_{t+1}) \right].$$

the value of $p_{n,t}$ is presented in the following proposition, the proof of which appears in the Appendix.

Proposition 1 *The log of the time t price of a zero-coupon bond with maturity n, $p_{n,t}$, can be written as:*

$$p_{n,t} = a_n + A_n^{'} Y_t, \tag{8}$$

where the scalar a_n and the $N \times 1$ vector A_n satisfy the following system of difference equations:

$$\begin{aligned} a_n &= a_{n-1} + \mu_m + \frac{1}{2}\Sigma_m^{'}\Sigma_m + (A_{n-1} - e_\pi)^{'}\left[\mu + \Sigma_H \Sigma_m\right] \\ &\quad + \frac{1}{2}(A_{n-1} - e_\pi)^{'} \Sigma_H \Sigma_H^{'} (A_{n-1} - e_\pi), \\ A_n^{'} &= \Gamma_m^{'} + \frac{1}{2}(\Sigma_{mf} \odot \Sigma_{mf})^{'} + (A_{n-1} - e_\pi)^{'}\left[A + \Sigma_{mf}^{'} \odot \Sigma_F\right] \\ &\quad + \frac{1}{2}\left[\Sigma_F^{'}(A_{n-1} - e_\pi) \odot \Sigma_F^{'}(A_{n-1} - e_\pi)\right]^{'}, \end{aligned} \tag{9}$$

with $a_0 = 0$, $A_0^{'} = (0, 0, ..., 0)$, and where e_π is an $N \times 1$ matrix with a 1 in the position that π_t occupies in the vector Y_t, and zeroes in all other positions.

Notice that the prices of all zero-coupon bonds (as well as their yields) take the form of affine functions of the state variables. Given the structure of Y_t, the term structure will represent a discrete-time multidimensional mixture of the Vasicek and CIR models. The process for the one-period short rate process, $r_t \equiv y_{1,t}$, is therefore simply $-(a_1 + A_1^{'} Y_t)$. Note that the pricing of real bonds (and the resulting real term structure of interest rates) is found by simply setting the vector e_π equal to a vector of zeroes.

Let $R_{n,t+1}^b$ and $r_{n,t+1}^b$ denote the nominal simple net return and log return, respectively, on an n-period zero coupon bond between dates t and $t+1$. Therefore:

$$\begin{aligned} R_{n,t+1}^b &= \exp(a_{n-1} - a_n + A_{n-1}^{'} Y_{t+1} - A_n^{'} Y_t) - 1, \\ r_{n,t+1}^b &= a_{n-1} - a_n + A_{n-1}^{'} Y_{t+1} - A_n^{'} Y_t. \end{aligned} \tag{10}$$

We now use the pricing model to value equity. Let V_t denote the real value of equity, which is a claim on the stream of real dividends, D_t. Using the real pricing kernel, V_t must satisfy the equation:

$$V_t = E_t\left[\exp(m_{t+1})(D_{t+1} + V_{t+1})\right]. \tag{11}$$

Using recursive substitution, the price-dividend ratio (which is the same in real or nominal terms), pd_t, can be written as:

$$pd_t = \frac{V_t}{D_t} = E_t\left\{\sum_{n=1}^{\infty} \exp\left[\sum_{j=1}^{n} (m_{t+j} + \Delta d_{t+j})\right]\right\}, \tag{12}$$

where we impose the transversality condition $\lim_{n\to\infty} E_t\left[\prod_{j=1}^{n} \exp(m_{t+j}) V_{t+n}\right] = 0$.

In the following proposition, we demonstrate that the equity price-dividend ratio can be written as the (infinite) sum of exponentials of an affine function of the state variables. The proof appears in the Appendix.

Proposition 2 *The equity price-dividend ratio, pd_t, can be written as:*

$$pd_t = \sum_{n=1}^{\infty} \exp\left(b_n + B_n^{'} Y_t\right), \tag{13}$$

where the scalar b_n and the $N \times 1$ vector B_n satisfy the following system of difference equations:

$$\begin{aligned} b_n &= b_{n-1} + \mu_m + \frac{1}{2}\Sigma_m^{'}\Sigma_m + (B_{n-1} + e_d)^{'}\left[\mu + \Sigma_H \Sigma_m\right] \\ &\quad + \frac{1}{2}(B_{n-1} + e_d)^{'} \Sigma_H \Sigma_H^{'} (B_{n-1} + e_d), \\ B_n^{'} &= \Gamma_m^{'} + \frac{1}{2}(\Sigma_{mf} \odot \Sigma_{mf})^{'} + (B_{n-1} + e_d)^{'}\left[A + \Sigma_{mf}^{'} \odot \Sigma_F\right] \\ &\quad + \frac{1}{2}\left[\Sigma_F^{'}(B_{n-1} + e_d) \odot \Sigma_F^{'}(B_{n-1} + e_d)\right]^{'}, \end{aligned} \tag{14}$$

with $b_0 = 0$, $B_0^{'} = (0, 0, ..., 0)$, and where e_d is an $N \times 1$ matrix with a 1 in the position that Δd_t occupies in the vector Y_t, and zeroes in all other positions. Given the expression for pd_t, the real value of equity can simply be written as $V_t = D_t \cdot pd_t$.

Comparing Equations (9) and (14), the stock price can be seen as the current dividend multiplied by the price of a "modified" consol bond. The "modified" consol bond has the following characteristics. First, the consol's coupons are real, and hence the inflation component characterized by the e_π term does not appear. Second, the payoffs each period are stochastic depending on how much dividends grow relative to D_t, hence the appearance of the e_d term.

Let R_{t+1}^s and r_{t+1}^s denote the nominal simple net return and log return, respectively, on equity between dates t and $t+1$. Therefore:

$$\begin{aligned} R_{t+1}^s &= \exp(\pi_{t+1} + \Delta d_{t+1}) \left(\frac{\sum_{n=1}^{\infty} \exp\left(b_n + B_n^{'} Y_{t+1}\right) + 1}{\sum_{n=1}^{\infty} \exp\left(b_n + B_n^{'} Y_t\right)} \right) - 1 \\ r_{t+1}^s &= (\pi_{t+1} + \Delta d_{t+1}) + \ln\left(\frac{\sum_{n=1}^{\infty} \exp\left(b_n + B_n^{'} Y_{t+1}\right) + 1}{\sum_{n=1}^{\infty} \exp\left(b_n + B_n^{'} Y_t\right)} \right). \end{aligned} \tag{15}$$

3 Examples of Affine Models

This Section provides a three-factor version of the general model developed in Section 2. We use this simplified framework to investigate two concrete examples within the three-factor model. In the first, the pricing kernel is exogenously specified; a feature quite common in the literature. In fact, we construct the economy to be as close in spirit to the modeling framework of Campbell and Shiller (1988) as possible. In the second, the pricing kernel is endogenous and reflects the intertemporal marginal rate of substitution of a representative agent with power utility, appended by slow-moving external habit as in Campbell and Cochrane (1999). This endogeneity not only imposes parameter restrictions across model equations and ties the parameters to structural parameters, but also restricts which stochastic processes enter the kernel as a function of the specification of the state variables in the economy. To show the tight link between exogenous discount rate models and models motivated by an equilibrium, we show how a particular version of the Campbell - Shiller economy can be interpreted as a particular parameterization of the representative agent endowment economy in Lucas (1978) and Mehra and Prescott (1985).

3.1 A Three-Factor Pricing Model

Here we specialize the general model presented in Section 2 to a three-factor model. Two of the factors are observable, fundamentals processes that are necessary to price nominal bonds and stocks: inflation π_t and dividend growth rates Δd_t. We will make inflation as neutral as possible in our examples, modeling it as a square root process with no effect on the real pricing kernel and independent of the other state variables. The implication of this modeling choice is that the Fisher hypothesis will hold in all our economies. The dividend process is richer and interacts with a third state variable. The third factor X_t is an unobservable state variable, such as a productivity shock or stochastic risk aversion. Whereas at this point the role and interpretation of this state variable remains unclear, in all the examples we will present it will provide us with the necessary degrees of freedom to match salient features of the short rate process.

The three-factor model is summarized as follows:

$$Y_t^{'} = (\Delta d_t, X_t, \pi_t) \tag{16}$$

$$m_{t+1} = \mu_m + \lambda_d \Delta d_t + \lambda_x X_t + s_d \varepsilon_{t+1}^d + \left[I \cdot s_x + (1-I) \cdot v_{mx} \sqrt{X_t} \right] \varepsilon_{t+1}^x$$

$$\begin{aligned}
\Delta d_{t+1} &= \mu_d + \rho_d \Delta d_t + g_x X_t + \sigma_d \varepsilon_{t+1}^d + \left[I \cdot \sigma_{dx} + (1-I) \cdot v_{dx} \sqrt{X_t} \right] \varepsilon_{t+1}^x \\
X_{t+1} &= \mu_x + g_d \Delta d_t + \rho_x X_t + \left[I \cdot \sigma_x + (1-I) \cdot v_x \sqrt{X_t} \right] \varepsilon_{t+1}^x \\
\pi_{t+1} &= \mu_\pi + \rho_\pi \pi_t + \sigma_\pi \sqrt{\pi_t} \varepsilon_{t+1}^\pi
\end{aligned}$$

$$\varepsilon_t^{'} = \left(\varepsilon_t^d, \varepsilon_t^x, \varepsilon_t^\pi\right)$$

where $\varepsilon_t \tilde{} N(\mathbf{0}, I)$, and where the indicator variable I takes the value 1 for the homoskedastic version of the model, and 0 for the heteroskedastic version of the model. The restricted VAR structure is necessary to ensure inflation neutrality. The potential heteroskedasticity in the model (if $I = 0$) is entirely driven by the X_t process. In the Appendix we demonstrate that this model falls within the general affine class. In addition, the three-factor model satisfies all of the restrictions listed in (6).

The solution for the pricing of bonds in the three-factor model is as follows:

$$p_{n,t} = A_n + B_n \Delta d_t + C_n X_t + D_n \pi_t, \tag{17}$$

where:

$$\begin{aligned} A_n &= A_{n-1} + \mu_m - \mu_\pi + B_{n-1}\mu_d + C_{n-1}\mu_x + D_{n-1}\mu_\pi \\ &\quad +1/2(s_d + B_{n-1}\sigma_d)^2 + I/2(s_x + B_{n-1}\sigma_{dx} + C_{n-1}\sigma_x)^2 \\ B_n &= \lambda_d + B_{n-1}\rho_d + C_{n-1}g_d \\ C_n &= \lambda_x + B_{n-1}g_x + C_{n-1}\rho_x + (1-I)/2(v_{mx} + B_{n-1}v_{dx} + C_{n-1}v_x)^2 \\ D_n &= -\rho_\pi + D_{n-1}\rho_\pi + 1/2(D_{n-1} - 1)^2\sigma_\pi^2, \end{aligned} \tag{18}$$

with $A_0 = B_0 = C_0 = D_0 = 0$. Note that all three factors, including dividend growth rates, are priced in the term structure. When $I = 1$, the model is a three factor Vasicek (1977) model, when $I = 0$, the model is a discrete-time version of a multifactor CIR model, with one factor following a square root process. The nominal interest rate equals:

$$\begin{aligned} r_t &= -(\mu_m - \mu_\pi + 1/2s_d^2 + I/2s_x^2) - \lambda_d \Delta d_t \\ &\quad - \left[\lambda_x + 1/2(1-I)v_{mx}^2\right] X_t + (\rho_\pi - 1/2\sigma_\pi^2)\pi_t, \end{aligned} \tag{19}$$

and the real rate of interest, $r_t^{\text{real}} = -\ln\{E_t[\exp(m_{t+1})]\}$, can be written as:

$$r_t^{\text{real}} = -(\mu_m + 1/2s_d^2 + I/2s_x^2) - \lambda_d \Delta d_t - \left[\lambda_x + 1/2(1-I)v_{mx}^2\right] X_t. \tag{20}$$

Note that the nominal short rate is equal to the sum of the real short rate and expected inflation, minus a constant term $(\sigma_\pi^2/2)$ due to Jensen's Inequality. The model thus yields an "approximate" version of the Fisher equation, where the approximation becomes more exact the lower the inflation volatility term.[7] In some examples below, we will parameterize the state variable dynamics such that the X_t process will be the real rate process.

The solution for the pricing of the equity price-dividend ratio in the three-factor model is as follows:

[7] The expected gross ex-post real return on a nominal one-period contract, $E_t[\exp(r_t - \pi_{t+1})]$ will be exactly equal to the gross ex-ante real rate, $\exp(r_t^{\text{real}})$.

$$pd_t = \sum_{n=1}^{\infty} \exp\left(a_n + b_n \Delta d_t + c_n X_t\right), \tag{21}$$

where:

$$\begin{aligned}
a_n &= a_{n-1} + \mu_m + (1 + b_{n-1})\mu_d + c_{n-1}\mu_x \\
&\quad +1/2\left[s_d + (1 + b_{n-1})\sigma_d\right]^2 + I/2\left[s_x + (1 + b_{n-1})\sigma_{dx} + c_{n-1}\sigma_x\right]^2 \\
b_n &= \lambda_d + (1 + b_{n-1})\rho_d + c_{n-1}g_d \\
c_n &= \lambda_x + (1 + b_{n-1})g_x + c_{n-1}\rho_x + (1 - I)/2\left[v_{mx} + (1 + b_{n-1})v_{dx} + c_{n-1}v_x\right]^2,
\end{aligned} \tag{22}$$

with $a_0 = b_0 = c_0 = 0$. Inflation neutrality implies that π_t is not priced in the price-dividend ratio.

We illustrate the intuition behind the general pricing equations (17)-(18) and (21)-(22) in the context of the example economies below. Here we introduce the risk premium on bonds and equities. Let $RP^b_{n,t}$ and $rp^b_{n,t}$ denote the nominal simple risk premium and nominal log risk premium, respectively, on an n-period zero coupon bond between dates t and $t+1$. These bond risk premiums can be expressed in closed-form:

$$\begin{aligned}
RP^b_{n,t} &\equiv E_t\left(R^b_{n,t+1}\right) - \exp(r_t), \\
&= \left\{\exp\left[-\left(\delta_n + \delta_{n,x}X_t + \delta_{n,\pi}\pi_t\right)\right] - 1\right\}\exp(r_t), \\
with &: \\
\delta_n &= B_{n-1}\sigma_d s_d + I\left(B_{n-1}\sigma_{dx} + C_{n-1}\sigma_x\right)s_x \\
\delta_{n,x} &= (1 - I)\left(B_{n-1}v_{dx} + C_{n-1}v_x\right)v_{mx} \\
\delta_{n,\pi} &= -D_{n-1}\sigma_\pi^2,
\end{aligned} \tag{23}$$

$$\begin{aligned}
rp^b_{n,t} &\equiv E_t\left(r^b_{n,t+1}\right) - r_t, \\
&= 1/2\left[s_d^2 + Is_x^2 - (s_d + B_{n-1}\sigma_d)^2 - I(s_x + B_{n-1}\sigma_{dx} + C_{n-1}\sigma_x)^2\right] \\
&\quad +1/2\left[(1 - I)v_{mx}^2 - (1 - I)(v_{mx} + B_{n-1}v_{dx} + C_{n-1}v_x)^2\right]X_t \\
&\quad +\left[(1 - D_{n-1}/2)D_{n-1}\sigma_\pi^2\right]\pi_t.
\end{aligned}$$

Note that Δd_{t+1} does not enter the expressions for the bond risk premiums since all of the heteroskedasticity in our model is driven by X_t and it is this heteroskedasticity that drives the time-variation in risk premiums.

Similarly, let RP^s_t and rp^s_t denote the nominal simple risk premium and nominal log risk premium, respectively, on equity. Thus, $RP^s_t \equiv E_t\left(R^s_{t+1}\right) - \exp(r_t)$, and $rp^s_t \equiv E_t\left(r^s_{t+1}\right) - r_t$. In some special cases, we will be able to provide closed-form solutions for the nominal simple equity risk premium. To do so, note that we can always express the conditional expected gross return on equity as:

$$E_t\left(1 + R^s_{t+1}\right) = E_t\left[\exp(\pi_{t+1} + \Delta d_{t+1})\left(\frac{pd_{t+1} + 1}{pd_t}\right)\right] \tag{24}$$

$$
\begin{aligned}
&= \frac{1}{pd_t} E_t \left[\sum_{n=1}^{\infty} \exp\left(a_n + (b_n+1)\Delta d_{t+1} + c_n X_{t+1} + \pi_{t+1}\right) \right] \\
&\quad + \frac{1}{pd_t} E_t \left[\exp\left(\Delta d_{t+1} + \pi_{t+1}\right) \right] \\
&= \frac{1}{pd_t} \left[\sum_{n=1}^{\infty} \exp\left(\bar{a}_n + \bar{b}_n \Delta d_t + \bar{c}_n X_t + \bar{d}_n \pi_t\right) \right] \\
&\quad + \frac{1}{pd_t} \exp\left(\hat{a}_n + \hat{b}_n \Delta d_t + \hat{c}_n X_t + \hat{d}_n \pi_t\right),
\end{aligned}
$$

with:

$$
\begin{aligned}
\bar{a}_n &= a_n + (b_n+1)\mu_d + c_n \mu_x + \mu_\pi + I/2[(b_n+1)\sigma_{dx} + c_n \sigma_x]^2 \\
\bar{b}_n &= (b_n+1)\rho_d + c_n g_d \\
\bar{c}_n &= (b_n+1) g_x + c_n \rho_x + (1-I)/2[(1+b_n) v_{dx} + c_n v_x]^2 \\
\bar{d}_n &= \rho_\pi + 1/2\sigma_\pi^2 \\
\hat{a}_n &= \mu_d + 1/2\sigma_d^2 + \mu_\pi + I/2\sigma_{dx}^2 \\
\hat{b}_n &= \rho_d \\
\hat{c}_n &= g_x + (1-I)/2 v_{dx}^2 \\
\hat{d}_n &= \bar{d}_n.
\end{aligned}
$$

3.2 An Extension of Campbell and Shiller (1988)

3.2.1 The Model

In Campbell and Shiller (1988), a linearized version of the present value model for pricing equities is developed. Their state variables are real dividend growth and a time-varying discount rate that they measure as the ex-post real return on commercial paper. Their state variables follow a VAR together with the log price-dividend ratio. The VAR is used to generate expectations of future state variables. Their model permits the testing of a present value model with constant expected excess returns (constant risk premium), along with a time-varying interest rate. This example is similar in spirit, but does not rely on linearization nor using a VAR to measure expectations. Our approach imposes more structure on the environment than does Campbell and Shiller (1988), since we fully specify the stochastic environment and generate price-dividend ratios that are an exact function of the state of the economy.

The unobserved state variable X_t in our version of the Campbell-Shiller present value framework will represent the real interest process. To accomplish this, we use equation (20), and set $\lambda_x = -1$, and $\lambda_d = s_x = v_{mx} = 0$. Furthermore, we do allow for a non-zero s_d and hence must put $\mu_m = -1/2s_d^2$. This model will fall within the homoskedastic class, and thus $I = 1$. In sum, the kernel process becomes:

$$Y_t^{'} = (\Delta d_t, X_t, \pi_t) \tag{25}$$
$$m_{t+1} = -1/2 s_d^2 - X_t + s_d \varepsilon_{t+1}^d$$

The state variables' dynamics are as in (16) with $I = 1$.

A process with no innovation ($s_d = 0$), would yield a viable pricing model where interest rates vary over time, but where the real return on equity has no risk premium. We assume only dividend innovations enter the kernel process. The correlation with dividends, and hence the parameter s_d, will determine the risk premium on equities and bonds. Assuming $s_x = 0$ implies that shocks to X_t are uncorrelated with shocks to the pricing kernel. Although we could have left s_x non-zero, it would have complicated the risk premium expressions and obscured the link with a the Mehra-Prescott equilibrium model that we document below.

3.2.2 Bond Pricing

The bond pricing equations for the Campbell and Shiller economy are determined by equations (17) through (20), with $I = 1$, $\lambda_x = -1$, $\mu_m = -1/2s_d^2$, and $\lambda_d = s_x = v_{mx} = 0$. Note, in particular, that the equation for the real rate of interest in (20) results in $r_t^{\text{real}} = X_t$.

By examining the nominal log risk premium, $rp_{n,t}^b$, displayed in (23), several important features of bond pricing in this model are brought out. All are a direct result of the homoskedasticity of the model, and not a result of any parameter restriction in this particular example. First, for a homoskedastic model ($I = 1$), $rp_{n,t}^b$ is equal to an affine function of inflation, where the function depends on the term of the bond. Thus, the nominal log risk premium for all bonds is unaffected by the current level of dividends and the real rate. The real log risk premium will be non-stochastic.

The parameter s_d plays an important role in determining the risk premium on bonds. Inspection of (23), for the case in which $I = 1$, reveals that $rp_{n,t}^b$ moves linearly with s_d. Specifically, $rp_{n,t}^b$ depends on s_d only through the constant $-s_d \sigma_d B_{n-1}$. Since the sign of B_{n-1} is ambiguous, the derivative of the bond risk premium with respect to s_d cannot be signed.

3.2.3 Stock Pricing

Imposing our parameter restrictions on equations (21) and (22), the terms b_n and c_n appear in a particularly simple form: $b_n = \rho_d + \rho_d b_{n-1} + g_d c_{n-1}$, and $c_n = g_x - 1 + g_x b_{n-1} + \rho_x c_{n-1}$.

The effects of changes in the real rate on the price-dividend ratio are captured by the c_n term. There are two effects of an increase in the real rate on the price of equity. There is a discount rate effect in which the price of equity decreases one-for-one with an increase in the real rate, and there is a cash flow effect in which the impact of the real rate changes on expected future cash flows is manifested in price changes. The cash flow effect is governed by the parameter g_x. If the real rate goes up by 1%, the conditional mean of dividend growth increases by g_x. The two effects are evidenced in the first two terms for c_n displayed in equation (37). It is possible that g_x can be negative, leading the two effects to both serve to decrease the price-dividend ratio and allowing for a greater variability in observed price-dividend ratios. It may in fact be economically reasonable for g_x to be negative, in which higher interest rates are accompanied by lower future expected cash flows.[8]

Under certain simplifying parameter restrictions, we can derive an explicit expression for the nominal simple risk premium on equity, RP_t^s, for this example. For $s_d = 0$, the equity risk premium has a simple expression:

$$RP_t^s = \left[\exp(\sigma_\pi^2 \pi_t) - 1\right] \exp(r_t), \qquad \text{for } s_d = 0. \tag{26}$$

The equity risk premium will be positive, and move with the short rate. However, the real equity risk premium will be precisely zero. This is intuitively clear, since with $s_d = 0$, the real dividend process is uncorrelated with the log of the pricing kernel, and thus represents nonsystematic risk. In such case, dividend risk would not be priced, and equities must yield a real expected return equal to the real rate of interest.

For the case in which $s_d \neq 0$, but in which $\rho_d = g_d = 0$, the equity risk premium can again be derived. In this case,

$$RP_t^s = \left[\exp(\sigma_\pi^2 \pi_t - s_d \sigma_d) - 1\right] \exp(r_t), \qquad \text{for } \rho_d = g_d = 0. \tag{27}$$

The equity risk premium will be decreasing in s_d. Thus, the parameter s_d controls the magnitude of both the bond and equity risk premia. Importantly, s_d plays a crucial role in determining both the bond and equity risk premia. It is indeed possible that

[8] Given the linear structure of the expressions for b_n and c_n, the difference equations can be solved analytically. Let

$$\phi \equiv \begin{pmatrix} \rho_d \\ g_x - 1 \end{pmatrix} \qquad W \equiv \begin{pmatrix} \rho_d & g_d \\ g_x & \rho_x \end{pmatrix}.$$

Then, it is straightforward to show that

$$\begin{pmatrix} b_n \\ c_n \end{pmatrix} = [I - W]^{-1} [I - W^n] \phi.$$

The limit effects (as $n \to \infty$) of a change in dividend growth or a change in the discount rate are thus:

$$\begin{pmatrix} b_\infty \\ c_\infty \end{pmatrix} = \begin{pmatrix} \frac{(1-\rho_x)\rho_d + g_d(g_x - 1)}{(1-\rho_x)(1-\rho_d) - g_x g_d} \\ \frac{g_x \rho_d + (1-\rho_d)(g_x - 1)}{(1-\rho_x)(1-\rho_d) - g_x g_d} \end{pmatrix}.$$

For example, when the cross-feedback effects are zero (the g terms), the ultimate impact of a change in the dividend growth rate on the price-dividend ratio equals $\frac{\rho_d}{1-\rho_d}$, and the ultimate impact of a change in the discount rate on the price-dividend ratio equals $\frac{-1}{1-\rho_x}$. Hence, if discount rates are much more persistent than cash flows, their impact on prices will be much larger.

s_d can have opposing effects on the two risk premia; while raising s_d will always lower the equity risk premium, it may raise the bond risk premium.

3.2.4 The Mehra-Prescott Model as a Special Case of the Campbell-Shiller Economy

Consider a simple equilibrium model in the tradition of Lucas (1978) and Mehra and Prescott (1985). We will see that this model can be placed into the framework of a restricted version of the Campbell-Shiller economy.[9]

A representative agent maximizes the expected discounted sum of a strictly increasing concave von Neumann-Morgenstern utility function U:

$$E_0 \left[\sum_{t=0}^{\infty} \beta^t U(C_t) \right], \tag{28}$$

where C_t is consumption at time t, β is a time discount factor, and E_t is the expectation operator conditional on all information up to time t.

In equilibrium, the consumption process C_t equals the exogenous aggregate real dividend process D_t. In addition, the first-order conditions of the optimization problem ensure that the following condition holds for all assets i and all time periods t:

$$1 = E_t \left[(1 + R_{i,t+1}) \frac{\beta U'(D_{t+1})}{U'(D_t)} \right], \tag{29}$$

where $R_{i,t+1}$ is the percentage real return on asset i over the period from t to $t+1$. Thus, as is well-known in this setting, the pricing kernel M_{t+1} is equivalent to the representative agent's intertemporal marginal rate of substitution.

We shall assume that the representative agent's utility function U has constant relative risk aversion equal to $\gamma > 0$, that is,

$$U(C) = \frac{C^{1-\gamma}}{1-\gamma}. \tag{30}$$

Therefore, we have:

$$M_{t+1} = \beta \left(\frac{D_{t+1}}{D_t} \right)^{-\gamma}. \tag{31}$$

The full description of the economy is completed with the specification of the dividend growth process and the inflation process. We shall assume that the dividend growth process is driven by a productivity shock, X_t:

[9] Bakshi and Chen (1997) develop a continuous-time version of the Lucas model that is closely related to our Mehra-Prescott economy. Labadie (1989) also adds stochastic inflation to a Mehra-Prescott economy, with a considerably different dividend process from ours.

$$\begin{aligned}
\Delta d_{t+1} &= \frac{\gamma}{2}\sigma_d^2 + \frac{\ln(\beta)}{\gamma} + \frac{X_t}{\gamma} + \sigma_d \varepsilon_{t+1}^d \\
X_{t+1} &= \mu_x + \rho_x X_t + \sigma_x \varepsilon_{t+1}^x \\
\pi_{t+1} &= \mu_\pi + \rho_\pi \pi_t + \sigma_\pi \sqrt{\pi_t}\varepsilon_{t+1}^\pi,
\end{aligned} \tag{32}$$

and therefore the log of the real kernel, m_{t+1}, equals $m_{t+1} = -\frac{\gamma^2}{2}\sigma_d^2 - X_t + \gamma\varepsilon_{t+1}^d$.

Specification (32) can be seen as a special case of the Campbell-Shiller specification. In particular, the Mehra-Prescott economy has two new parameters, β and γ, and imposes the following parameter restrictions:

$$\begin{aligned}
s_d &= -\sigma_d \gamma \\
\mu_d &= 1/2\gamma\sigma_d^2 + \ln(\beta)/\gamma \\
g_x &= 1/\gamma \\
\sigma_{dx} &= \rho_d = g_d = 0.
\end{aligned} \tag{33}$$

This model economizes on parameters in two ways. First, it simplifies the joint dynamics of dividend growth and the real rate. The dividend growth process is now driven by a productivity shock X_t, but has no direct autoregressive component. The real rate/productivity shock follows an autoregressive process. Second, cross-equation parameter restrictions are imposed as the kernel process is fundamentally related to the dividend growth process.

In the empirical work that follows, we will make this benchmark economy a bit more interesting by making the real rate a square root process:

$$X_{t+1} = \mu_x + \rho_x X_t + v_x \sqrt{X_t}\varepsilon_{t+1}^x. \tag{34}$$

This implies that the nominal rate of interest, r_t, can be written as:

$$r_t = \mu_\pi + X_t + \left(\rho_\pi - 1/2\sigma_\pi^2\right)\pi_t. \tag{35}$$

Whereas the real rate follows a one-factor CIR model, the nominal rate follows a two-factor model that is similar to the continuous-time model of Richard (1978).

The nominal simple risk premium and nominal log risk premium on an n-period zero coupon bond, $RP_{n,t}^b$ and $rp_{n,t}^b$ respectively, can be written as:

$$\begin{aligned}
RP_{n,t}^b &= \left[\exp\left(\sigma_\pi^2 D_{n-1}\pi_t\right) - 1\right]\exp(r_t), \\
rp_{n,t}^b &= -\frac{1}{2}v_x^2 C_{n-1}^2 X_t + \sigma_\pi^2 D_{n-1}\left(1 - \frac{1}{2}D_{n-1}\right)\pi_t.
\end{aligned} \tag{36}$$

Notably, the real simple bond risk premium is equal to zero, and the real log bond risk premium is equal to $-\frac{1}{2}\sigma_x^2 B_{n-1}^2 X_t$, which is proportional to the current real rate of interest.

The pricing of equities is greatly simplified in this framework. The expression for c_n becomes:

$$c_n = -1 + \frac{1}{\gamma} + \rho_x c_{n-1} + 1/2 v_x^2 c_{n-1}^2. \tag{37}$$

If the representative agent is more risk averse than an investor with log utility ($\gamma > 1$), then $\frac{\partial pd_t}{\partial X_t} = \sum_{n=1}^{\infty} c_n \cdot \exp(a_n + c_n X_t)$ will generally be negative, and thus increases in the real rate will lower the price-dividend ratio. This is because c_n will generally be negative when $\gamma > 1$, as long as v_x^2 is small relative to ρ_x. The intuition for this result is simple. There are two competing effects of a change in the real rate. First, an increase in the real rate leads to an increase in the expected return on all assets, leading to a fall in the price of equity. Second, in this example an increase in the real rate (the technology shock) also raises the conditional mean of dividend growth, and hence, leads to an increase in the price of equity. From equation (32), a 1% increase in X leads to a $\frac{1}{\gamma}$% increase in $E_t(\Delta d_{t+1})$. This combination of forces is apparent in the first two terms in the expression for c_n in equation (37). The -1 term reflects the discount rate effect, and the $\frac{1}{\gamma}$ term reflects the cash flow effect.[10] For $\gamma > 1$,the discount rate effect dominates the cash flow effect, and the degree of domination will be greater the larger is γ. The parameter restrictions the equilibrium economy imposes on the Campbell-Shiller economy will likely have severe consequences for its empirical performance. The fact that there are two competing effects on the price-dividend ratio will lead to a general lack of variability in the price-dividend ratio.

It is now also straightforward to derive a closed-form expression for the equity premium, since it is a special case of Equation (27) above with $s_d = -\sigma_d \gamma$. Hence,

$$RP_t^s = \left[\exp(\gamma\sigma_d^2 + \sigma_\pi^2 \pi_t) - 1\right] \cdot \exp(r_t). \tag{38}$$

As would be expected, the risk premium is increasing in the degree of risk aversion and the volatility of dividend growth.

3.3 The "Moody" Investor Economy

3.3.1 The Model

Consider an economy as in Lucas (1978), but modify the preferences of the representative agent to have the form:

$$E_0 \left[\sum_{t=0}^{\infty} \beta^t \frac{(C_t - H_t)^{1-\gamma} - 1}{1-\gamma} \right], \tag{39}$$

where C_t is aggregate consumption and H_t is an exogenous "external habit stock" with $C_t \geq H_t$.

[10]The remaining terms account for further effects due to the persistence in X_t and a Jensen's inequality term.

One motivation for an "external" habit stock is the framework of Abel (1990, 1999) who specifies preferences where H_t represents past or current aggregate consumption, which a small individual investor takes as given, and then evaluates his own utility relative to that benchmark. That is, utility has a "keeping up with the Joneses" feature. In Campbell and Cochrane (1999), H_t is taken as an exogenously modelled subsistence or habit level. Hence, the coefficient of relative risk aversion equals $\gamma \cdot \frac{C_t}{C_t - H_t}$, where $\left(\frac{C_t - H_t}{C_t}\right)$ is defined as the surplus ratio. As the surplus ratio goes to zero, the consumer's risk aversion goes to infinity. In our model, we view the inverse of the surplus ratio as a preference shock, which we denote by Q_t. Thus, $Q_t = \frac{C_t}{C_t - H_t}$. Risk aversion is now characterized by $\gamma \cdot Q_t$, and $Q_t > 1$.

The marginal rate of substitution in this model determines the real pricing kernel. Taking the ratio of marginal utilities of time $t+1$ and t and imposing $C_t = D_t$, we obtain:

$$\begin{aligned} M_{t+1} &= \beta \frac{(C_{t+1}/C_t)^{-\gamma}}{(Q_{t+1}/Q_t)^{-\gamma}} \\ &= \beta \exp\left[-\gamma \Delta d_{t+1} + \gamma \left(X_{t+1} - X_t\right)\right], \end{aligned} \tag{40}$$

where $X_t = \ln(Q_t)$.

This model may better explain the predictability evidence than the Mehra-Prescott model. The evidence suggests that expected returns and the price of risk move countercyclically. Using the intuition of Hansen-Jagannathan (1991) bounds, we know that the coefficient of variation of the pricing kernel equals the maximum Sharpe ratio attainable with the available assets. As Campbell and Cochrane (1999) also note, with a log-normal kernel:

$$\frac{\sigma_t(M_{t+1})}{E_t(M_{t+1})} = \sqrt{\exp\left[Var_t\left(m_{t+1}\right) - 1\right]}. \tag{41}$$

Hence, the maximum Sharpe ratio characterizing the assets in the economy is an increasing function of the conditional volatility of the pricing kernel. If we can construct an economy in which the conditional variability of the kernel varies through time and is higher when Q_t is high (that is, when consumption has decreased closer to the habit level), then we have introduced the required countercyclical variation into the price of risk. Note that our previous models fail to accomplish this. The conditional variability of the nominal pricing kernel in the Campbell-Shiller economy depends on the level of inflation, which tends to move pro-cyclically, and the conditional variability of the real kernel was constant.

Whereas Campbell and Cochrane (1999) have only one source of uncertainty, namely, consumption growth, which is modeled as an i.i.d. process, we embed the "Moody Investor" economy in our three-factor model discussed in Section 3.1. The unobserved state variable, X_t, now is stochastic risk aversion. Although the intertemporal marginal rate of substitution determines the form of the real pricing kernel through (40), we still have a choice on how to model Δd_t and X_t. Since $Q_t > 1$, we model $X_t \equiv \ln(Q_t)$ as a square-root process. To capture the notion that stochastic risk aversion behaves countercyclically, we allow the shocks to dividend growth and stochastic risk aversion to be correlated, expecting this correlation to be negative. That is, a positive shock to dividend growth is expected to reduce risk aversion as it leads to an increase in the surplus ratio. We also allow for autocorrelation in the dividend process.

To summarize, the economy is as follows:

$$\begin{aligned} m_{t+1} &= \ln(\beta) - \gamma \cdot \Delta d_{t+1} + \gamma \cdot (X_{t+1} - X_t) \\ &= \ln(\beta) - \gamma\left(\mu_d - \mu_x\right) - \gamma\rho_d \Delta d_t + \gamma(\rho_x - 1)X_t - \gamma\sigma_d\varepsilon^d_{t+1} + \gamma\left(v_x - v_{dx}\right)\sqrt{X_t}\varepsilon^x_{t+1} \end{aligned} \tag{42}$$

$$\begin{aligned} \Delta d_{t+1} &= \mu_d + \rho_d \Delta d_t + \sigma_d \varepsilon^d_{t+1} + v_{dx}\sqrt{X_t}\varepsilon^x_{t+1} \\ X_{t+1} &= \mu_x + \rho_x X_t + v_x \sqrt{X_t}\varepsilon^x_{t+1} \\ \pi_{t+1} &= \mu_\pi + \rho_\pi \pi_t + \sigma_\pi \sqrt{\pi_t}\varepsilon^\pi_{t+1} \end{aligned}$$

This model easily fits into the three-factor structure, with the parameter restrictions $\mu_m = \ln(\beta) - \gamma\left(\mu_d - \mu_x\right)$, $\lambda_d = -\gamma\rho_d$, $\lambda_x = \gamma(\rho_x - 1)$, $s_d = -\gamma\sigma_d$, $v_{mx} = \gamma\left(v_x - v_{dx}\right)$, and $g_x = g_d = 0$.

The real kernel process, m_{t+1}, is heteroskedastic ($I = 0$), with its conditional variance moving with X_t. In particular,

$$Var_t(m_{t+1}) = \gamma^2\sigma_d^2 + \gamma^2 \cdot \left(v_x - v_{dx}\right)^2 X_t. \tag{43}$$

Consequently, increases in X_t will increase the Sharpe Ratio of all assets in the economy, and the effect will be greater the larger are γ, v_x, and $|v_{dx}|$. If X_t and Δd_t are negatively correlated, with $v_{dx} < 0$, the Sharpe Ratio of assets will increase during economic downturns (falls in Δd_t).

3.3.2 Bond Pricing

The bond pricing equations for the Moody Investor economy are determined by equations (17) through (20), noting the specifications $I = 0$, $\mu_m = \ln(\beta) - \gamma\left(\mu_d - \mu_x\right)$, $\lambda_d = -\gamma\rho_d$, $\lambda_x = \gamma(\rho_x - 1)$, $s_d = -\gamma\sigma_d$, and $v_{mx} = \gamma\left(v_x - v_{dx}\right)$. Of particular interest is the resulting real rate of interest in the Moody Investor economy, $r_t^{\text{real,M-I}}$, which can be written as:

$$r_t^{\text{real,M-I}} = -\ln(\beta) + \gamma\left(\mu_d - \mu_x\right) - 1/2\gamma^2\sigma_d^2 + \gamma\rho_d\Delta d_t + \left[\gamma\left(1 - \rho_x\right) - 1/2\gamma^2\left(v_x - v_{dx}\right)^2\right] X_t. \tag{44}$$

In this model we did not parameterize the dynamics of the state variables so as to yield a simplified interest rate process. Now, the interest rate is totally endogenous and a function of the dividend growth rate and stochastic risk aversion dynamics. This implies that this economy may be subject to the low risk-free rate puzzle [Kocherlakota (1996), Weil (1989)].

To understand the risk-free rate in equation (44), first consider the risk-free rate in the standard Mehra-Prescott economy, $r_t^{\text{real,M-P}}$:

$$r_t^{\text{real,M-P}} = -\ln(\beta) + \gamma E_t(\Delta d_{t+1}) - \frac{1}{2}\gamma^2 V_t(\Delta d_{t+1}). \tag{45}$$

The first term represents the impact of the discount factor. The second term represents a consumption-smoothing effect. Since in a growing economy agents with concave utility ($\gamma > 0$) wish to smooth their consumption stream, they would like to borrow and consume now. This desire is greater, the larger is γ. Thus, since it is typically necessary in Mehra-Prescott economies to allow for large γ to generate a high equity premium, there will also be a resulting real rate that is higher than empirically observed. The third term is the standard precautionary savings effect. Uncertainty induces agents to save, therefore depressing interest rates and mitigating the consumption-smoothing effect.

The real rate in the Moody investor economy, $r_t^{\text{real,M-I}}$, equals the real rate in the Mehra-Prescott economy, plus two additional terms:

$$r_t^{\text{real,M-I}} = r_t^{\text{real,M-P}} + \gamma\left[(1-\rho_x)X_t - \mu_x\right] - \frac{1}{2}\gamma^2\left(v_x^2 - 2v_x v_{dx}\right)X_t. \tag{46}$$

The first of the two extra terms represents an additional consumption-smoothing effect. In this economy, risk aversion is also effected by X_t, and not only γ. When X_t is above its unconditional mean, $\mu_x/(1-\rho_x)$, the consumption-smoothing effect is exacerbated. The second of the two extra terms represents an additional precautionary savings effect. The uncertainty in stochastic risk aversion has to be hedged as well. Since v_{dx} is expected to be negative, this hedging term is negative as well.

3.3.3 Stock Pricing

The stock pricing equations for the Moody Investor economy are determined by equations (21) and (22). In this economy, the terms b_n and c_n can be written as:

$$\begin{aligned} b_n &= -\gamma\rho_d + \rho_d + \rho_d b_{n-1} \\ c_n &= -\gamma(1-\rho_x) + \rho_x c_{n-1} + 1/2\left[(1+b_{n-1}-\gamma)v_{dx} + (c_{n-1}+\gamma)v_x\right]^2. \end{aligned} \tag{47}$$

This model provides an alternative to the preference-free Campbell-Shiller framework for breaking the tight link between cash flow and discount rate effects. It is still the case that a shock that decreases the dividend growth rate simultaneously depresses cash flows and discount rates, which have countervailing effects on prices.

However, there is now an additional discount rate effect that makes the cash flow effect more pronounced. Since X_t and Δd_t are negatively correlated, a negative shock to dividend growth (recession) leads to higher risk aversion. Higher risk aversion serves to lower prices and the price-dividend ratio. These effects can be seen in the expressions for the b_n and c_n coefficients in equation (47). The direct discount rate effect is represented by the $-\gamma \cdot \rho_d$ term in b_n, hence when dividend growth decreases, prices increase by $\gamma \cdot \rho_d$. From the dynamics of Δd_t, the direct cash flow effect would be a decrease in the price-dividend ratio of ρ_d. The direct effect of the resulting positive shock to X_t is represented by the $-\gamma (1 - \rho_x)$ term in c_n. Thus, prices are further depressed by $\gamma (1 - \rho_x)$. The other coefficients accommodate the persistence in the process.

4 Estimation and Asset Return Properties

In this section, we begin by outlining the general estimation methodology for the model parameters. We then briefly discuss the data. Next, we discuss the qualitative properties of the parameter estimates. Finally, we analyze the implied unconditional moments of bond and stock returns under each example economy, and compare them with those estimated from the data. Although our main examples are the Campbell-Shiller and Moody Investor economies, we also include the Mehra-Prescott economy as an easily recognizable benchmark.

4.1 General Methodology

The three example economies above have a very similar structure. In particular, the two "measurable" economic factors in all three economies are inflation and dividend growth. Moreover, all three economies have one state variable, X_t, that we do not directly measure from the data: the real rate in the Campbell-Shiller and Mehra-Prescott economies, and the stochastic risk aversion process in the Moody Investor economy. The state variable vector for our three economies is $Y_t^{'} = [\Delta d_t, X_t, \pi_t]$. Now consider the vector $W_t^{'} = [\Delta d_t, r_t, \pi_t]$, recalling that r_t represents the nominal interest rate. Let the parameters governing the state variables and pricing kernel be represented by the vector Ψ. The affine structure implies:

$$W_t = c(\Psi) + C(\Psi) Y_t, \tag{48}$$

where $c(\Psi)$ is a 3×1 vector and $C(\Psi)$ a 3×3 matrix of structural coefficients. Using the stochastic process describing the dynamics of Y_t, it is straightforward to derive a structural VAR relation for W_t:

$$W_t = d(\Psi) + D(\Psi) W_{t-1} + C(\Psi) \left(\Sigma_F F_{t-1} + \Sigma_H \right) \varepsilon_t, \tag{49}$$

where it is understood that the change in variables from Y_t to W_t is made in F_{t-1} as well, and:

$$\begin{aligned} d(\Psi) &= C(\Psi)\mu + \left(I - C(\Psi)A\left[C(\Psi)\right]^{-1}\right)c(\Psi), \\ D(\Psi) &= C(\Psi)A\left[C(\Psi)\right]^{-1}. \end{aligned} \tag{50}$$

Since ε_t was assumed to be normally distributed with identity covariance matrix, maximum likelihood estimation is one possibility to obtain estimates of Ψ. For reasons that will soon become clear, we will use standard GMM. Given the relation between W_t and Y_t in Equation (48), computation of the moments of Y_t leads immediately to the moments of W_t. We will restrict attention to the first two moments (given the log-normal structure). In particular,

$$\begin{aligned} E(Y_t) &= (I - A)^{-1}\mu, \\ vec[Var(Y_t)] &= (I - A \otimes A)^{-1} vec\left[\Sigma_F\left(E\left[Y_t\right] \odot I\right)\Sigma_F' + \Sigma_H \Sigma_H'\right], \\ cov(Y_t Y_{t-1}') &= A\, Var(Y_t). \end{aligned} \tag{51}$$

These moments ignore the presence of the function $\shortparallel \cdot \shortparallel$ in Equations (1) - (2). Although it is possible to derive the exact relations, they will not dramatically alter our results as long as the mass below zero is small.

Of the examples in Section 3, the Mehra-Prescott economy is the most parsimonious; it has 9 parameters whereas the Campbell-Shiller economy has 13 and the Moody Investor economy has 12. As a consequence, it is possible to identify all the parameters from the first and second moments of $W(t)$, for example in an exactly identified GMM system. Such an approach would then match some moments of the nominal rate process exactly. Instead, we choose to work with economies that match salient properties of dividend growth, inflation, the nominal short rate and that match the equity risk premium. To accomplish this, we fix the critical parameter (γ in the utility-based models and s_d in the Campbell-Shiller model) in order to match the equity premium (measured in logs) in the data and then re-estimate the remaining parameters. Specifically, this involves using our analytical solutions to set a reasonable initial estimate for the risk parameter, estimating the other parameter values using the GMM and re-iterating a few times until an endogenous risk premium mean results that is close to that in the data.[11] As we will further discuss below, matching the equity premium in our models is not as big a challenge as is portrayed in the equity premium literature, since we use dividend and not consumption data.

With the various economies fully parameterized, we can use Equation (48) to recover the state variables relevant for our particular observed sample. We then investigate other return properties predicted by the model by computing small sample

[11] Alternatively, we could have appended equity moments to the set of moments to match, but since these moments involve infinite sums the computational time is much increased.

moments for model-implied bond and stock returns. Hence our empirical procedure does not use population moments and thus avoids small sample problems in statistical inference. Given the obvious stochastic singularities in all of the models, it would not be very hard to reject them. However, it remains useful to test and examine which moments the model can and cannot match.

4.2 Data Properties

The data inputs for this paper are annual stock and bond returns, a one-year nominal rate, inflation, a long-term bond yield, and dividend growth rates, all for the U.S. Most of the data are from the Ibbotson Database. We use annualized data to avoid the seasonality in dividend payments. Both Campbell and Shiller (1988) and Cochrane (1992) use annual data for this reason. The use of annual data also diminishes the small mis-matches that occur, for example, in matching inflation data collected during the month with asset price data.

To arrive at an annual dividend to be used in computing dividend growth rates and dividend yields, Campbell and Shiller simply add the dividends paid out during the year, whereas Cochrane measures the aggregate dividends assuming they were invested in the market. Below, we will show the properties of dividend growth rates using both of these assumptions. For stock returns, we use the actual total returns with re-invested dividends.

The one-year interest rate is supplied by Ibbotson. It represents the yield on Treasury bills with maturity closest to one year. The Ibbotson bond series uses a one bond portfolio with a term of approximately 20 years. We define the yield on this bond series as our long rate. Unfortunately, the yield data series only goes back to the 1950's. We obtain a time series of the yield on a similar bond portfolio from statistics supplied by the Board of Governors. For the overlapping years, the correlation between the two series is 97%.

Table 1 indicates the data sources for the time series we use and their availability. In Table 2 we analyze the time-series properties of the "state variables," the exogenous variables in our model and the "instruments," the variables that are most typically used to empirically track predictable components in returns. These instruments include the term spread, dividend yield and nominal interest rate. In our model, the instruments are endogenous. Note that the instruments and state variables are mostly quite persistent time series, except for real dividend growth. Long-term bonds on average yield about 1% more than a one-year bond investment.

4.3 Parameter Estimates

In this section we discuss the qualitative properties of the estimation results and some important parameter values.

4.3.1 Mehra-Prescott Economy

As indicated above, we begin by fixing γ at reasonable values ranging from 2 to 10 and re-estimate the 8 remaining structural parameters to match 9 moments: the mean, variance and autocovariance of dividend growth, inflation, and the nominal rate of interest. Since there is one over-identifying restriction, we can use the standard J-test [Hansen (1982)] to verify that the fit with the moments is satisfactory. We did not reject the restrictions for any of the parameters we tried.

Since the economy is fully parameterized, we can recover the implied state variables for the sample using equation (48). The implied real rate process is persistent with an autocorrelation of about 0.85 (with a standard error of 0.33). Table 3 reports the implied mean excess equity return, at different levels of γ.[12] The last line in the table reports the corresponding data moment with a GMM standard error, revealing the equity premium (in logs) to be 6.14% with a standard error of 2.40. As γ increases, agents with greater risk aversion will value the exogenous dividend stream less and require higher expected returns in order to hold the claim to it.

Whereas in the original Mehra-Prescott paper, the equity premium could not be matched for moderate levels of risk aversion, our economy produces an equity premium larger than that in the data at $\gamma = 6$. The main reason is the use of dividend growth rates as the fundamental process. The variability of dividend growth is an order of magnitude larger than the variability of consumption growth (see Table 2) and as Equation (38) shows, the equity risk premium is directly impacted by the product of dividend growth variability and γ. At $\gamma = 5.55$, we obtain a value for the equity premium that is indistinguishable from that of the data.[13] We use that risk aversion value and the corresponding parameter estimates for the remainder of the analysis.

4.3.2 Campbell-Shiller Economy

For the Campbell-Shiller model, we use 12 moments to estimate all parameters except s_d. The moments we add to the ones used in the Mehra-Prescott world are cross-moments between dividend growth and interest rates. After some experimentation, we find that the equity premium is almost matched at a s_d value of -0.45.[14] Another

[12] It is straightforward to compute population moments as well, either analytically (as in the case of bond variables) or by simulation (as in the case of equity variables).

[13] Hagiwara and Herce (1997) also construct an asset pricing model using dividends. They are able to explain the equity premium with a much lower level of risk aversion than in models using consumption data. Burnside (1994) uses both consumption and dividend growth rates to price equities.

[14] Recall that for $\rho_d = g_d = 0$, the two models are linked by $s_d = -\gamma \cdot \sigma_d$. With σ_d estimated to be 0.114 in the Mehra-Prescott estimation and γ equal to 5.55, we deduce that s_d's in the range of -0.5 to -0.7 should be tried. In reality ρ_d and g_d are not zero, ρ_d is estimated to be 0.175 with a standard error of 0.185, and g_d is estimated to be -0.035 with a standard error of 0.012. Hence, these parameters are indeed either statistically insignificant from zero, or close to zero in magnitude,

parameter of interest is g_x. In the Mehra-Prescott model g_x was constrained by the structure of the model to equal $1/\gamma$; here it is a free parameter. We estimate g_x to be -0.605, but it is not estimated very precisely (the standard error is 2.33). That implies that at the estimated value, this model will likely generate more price-dividend variability, since shocks to real rates now generate cash flow and discount rate effects in the same direction (an increase in X_t increases the discount rate, depressing prices, but also depresses cash flows, which in turn depresses prices further). We will discuss the magnitude of this effect below.

4.3.3 Moody Investor Economy

For the Moody Investor Economy, we attempt to estimate all parameters from the same 12 moments that were used for the Campbell-Shiller model. For γ equal to 2.6, we obtain an equity premium very close to the one observed in the data.[15] The parameters of interest here are of course the ones driving the X_t process, since these determine how much variation there will be in risk aversion and hence the price of risk. With standard errors between parentheses, the estimates were 0.233 (0.085) for $\mu_x/(1-\rho_x)$, 0.358 (0.100) for ρ_x and 0.099 (0.114) for v_x.

What are the implications of these parameter estimates? First recall that risk aversion is stochastic in this model, equalling γQ_t. Hence, our parameters imply a time series of risk aversion. The average risk aversion coefficient is 3.29 and its standard deviation over the sample is only 0.17. Very high risk aversion is not required as in Campbell and Cochrane (1999) because of the higher variability of dividend growth. Second, risk aversion is indeed positively correlated with recessions, and reaches its peak in the Great Depression, while still remaining below 4.0. One interpretation of this behavior of risk aversion, and hence the price of risk in this model, is the wealth-based risk premium idea of Sharpe (1990). Sharpe postulates that when people become wealthier their risk aversion drops. This has only price implications when it happens for society as a whole, that is, when aggregate economic growth has been unusually high propelling wealth levels above normal levels. Third, does the relation between X_t and current and past dividend levels conform to a habit formation story? It is straightforward, using the same first-order approximation as Campbell and Cochrane, to write the log habit level as slowly decaying moving average of past

explaining why our guess for the relevant range was rather accurate.

[15]The estimation for this model was decidedly less smooth, and we had trouble obtaining convergence, for example because autocorrelation parameters drifted into non-stationary regions. We finally dropped two cross-moments and fixed the parameter for the unconditional mean of dividend growth at its sample value. That yielded an exactly identified system for which reasonable parameter values were obtained, but with huge standard errors. These are not so surprising since in this model the parameters v_{dx} and v_x are hard to identify jointly. In our model, the critical sensitivity ratio, determining how risk aversion reacts to dividend shocks, equals $-v_{dx}/v_x$. In Campbell and Cochrane (1999), this sensitivity ratio is explicitly modelled as a time-varying, non-linear process. We fixed v_{dx} at its estimated value, as expected smaller than 0 (-0.246), and re-estimated, now obtaining more reasonable standard errors.

consumption, but the relation is more complex because of the presence of separate X-shocks and the autocorrelation in consumption growth. However, Campbell and Cochrane have more flexibility in modelling the sensitivity of the surplus ratio to consumption shocks and they ensure that the derivative of log(habit) with respect to log(consumption) is always positive. In our model, this condition corresponds to $v_x/v_{dx} \geq 1 - \exp(X_{t+1})$ for all t. Although we could impose parameter restrictions that would make this condition likely to hold, we choose to let the data "speak." At the current parameter values, this particular restriction is not satisfied, but it would be if we were to drop v_{dx} to $-.40$.

4.4 Unconditional Properties of Asset Returns

Table 4 reports the mean price-dividend ratio and the mean and variance of the equity premium (in logs) for all three models. By construction, the equity premium is matched by all three economies. The mean price-dividend ratio is similar across the three examples at around 16.4, which is substantially lower than what is observed in the data, where it is 25.23. One potential explanation is that the price-dividend ratio mean in the data is upwardly biased, because of the recent trend of distributing cash to shareholders through repurchases rather than dividends, but using the data in Cole, Helwege and Laster (1996), such an adjustment is of much smaller magnitude. Most noticeable about the table, is that the additional richness of the Campbell and Shiller and Moody Investor economies leads to higher, and more realistic variability of equity returns.

In Table 5 we look at the term structure implications of the three models. In the data, we observe on average a positive term spread and bond premium (in logs). Also, we observe a bond return volatility of about 8%, which is much lower than the equity return variability, and low correlation between equity and bond returns. In the Mehra-Prescott model, although excess bond return variability is of the right order of magnitude compared to the data, the model generates an on average downward-sloping yield curve and a negative bond premium. Both other models, however, do generate positive bond premiums and they also generate more variability in excess bond returns. Hence, we need the added flexibility of the Campbell-Shiller type economy or the Moody Investor economy to be able to simultaneously match simple mean and variance properties of both bond and equity returns.

A final interesting statistic to examine is the correlation between bond and equity returns. As the final column in Table 5 indicates, the correlation in the data is quite low (0.189), although it is not very precisely estimated. Despite equity behaving as a consol bond in all of our economies, the role of stochastic dividends is such that the endogenous correlation between bond returns and equity returns need not be very high. In particular, it is even slightly negative for the Mehra-Prescott economy, but all economies generate correlations within a two-standard error band of the empirical estimate. The Moody Investor economy comes closest to matching the data moment with a bond/equity correlation of 0.10.

In Table 6, we examine whether these example models can replicate the non-linearities in stock and bond returns. In the data, both equity and bond returns display leptokurtosis, but equity returns are negatively skewed whereas bond returns are positively skewed. All models generate negatively skewed equity returns as in the data, but produce too much kurtosis in equity returns. They also match the positive skewness and excess kurtosis in bond returns. It is important to remember that these are small sample results. For example, log bond returns in the Campbell-Shiller economy should be normally distributed, since we did not allow for heteroskedasticity in the state variables. The skewness and kurtosis we see here are purely a small sample phenomenon (as they may be in the real data).[16]

5 Empirical Analysis of Predictability

This section examines the performance of the various models with respect to predictability, using a variety of measures. We compute variance ratios to measure long-run autocorrelations in returns, we estimate univariate predictability regressions with "yield" variables, we analyze empirical and model-based conditional risk premiums, and finally, we compute the variability of price-dividend ratios.

5.1 Variance Ratios

In Table 7 we report variance ratios for both stock and bond returns. The variance ratios observed in the data suggest some long-run persistence in bond returns (variance ratios above 1), whereas the evidence for stocks suggests some slight mean reversion (variance ratios below 1), consistent with the well-known evidence in Poterba and Summers (1988). The Mehra-Prescott economy generates slightly too much persistence in stock returns, and too much mean reversion in bond returns, with the latter being significantly different from the positive sample variance ratios. In the Campbell and Shiller world, the relative magnitudes are more realistic, in that equity returns are much more mean-reverting than bond returns, but the model also fails to generate positive persistence in bond returns. The Moody Investor economy is the only one that generates some weak positive persistence in bond returns, and strong mean reversion in equity returns.

[16] As a check, we compute kurtosis and skewness for our three example economies in population, using a simulation of 25,000 observations. As expected, for the Campbell-Shiller economy, we indeed find a normal population distribution for log bond returns as well as for equity returns. Interestingly, the dramatic excess kurtosis and negative skewness for equity returns generated by the Mehra-Prescott economy are not present in the population moments. Hence, the negative skewness observed in equity returns can be matched in an economy which in population generates symmetric equity returns. The extreme realizations of dividend growth during the Depression years are the likely cause of this phenomenon.

It is important to realize that variance ratios are biased downward in small samples and that the asymptotic standard errors we use to compare data with model moments may not be appropriate in our small sample. Nevertheless, both are based on the same small sample and hence the bias in both computations may be similar.[17]

5.2 Univariate Predictability Regressions

To examine linear predictability, we focus on two measures of "yield" as predictive instruments: the dividend yield in excess of the nominal interest rate [see Harvey (1991)], and the long-term yield in excess of the short rate (the term spread).[18]

The univariate regressions in the data, reported in the last row of Table 8, provide weak evidence of predictability in the stock return equation. Both an increase in the term spread and the dividend yield indicate a higher risk premium on equity. Whereas the excess dividend yield fails to predict the future stock return significantly at the 10% level, the term spread coefficient is significantly different from zero at the 10%, but not the 5% level. The sign and magnitude of the coefficients are similar to the coefficients found in previous studies. One reason for the weak predictability results is the annual data frequency, as most predictability studies use monthly data. In addition, the literature has typically found stronger evidence of predictability for longer-horizon returns.

The univariate bond return regressions reveal that the dividend yield does not seem to predict bond returns. However, the coefficient is positive, as it was in the stock return equation. The term spread is a very strong predictor of excess bond returns. This result is very closely related to one of the long-standing puzzles in the term structure literature. Campbell and Shiller (1991) point out that the yield spread provides the wrong prediction for changes in future long rates relative to the prediction implied by the Expectations Hypothesis. In particular, when one regresses the change in the long rate multiplied by the duration of the bond onto a constant and the yield spread, one finds significantly negative coefficients that become more negative for longer maturities. Changing signs in the regression and adding the yield spread, the dependent variable becomes an excess bond return. The regression coefficient that we find is then approximately one minus the regression coefficient in the Campbell-Shiller

[17] To examine the effect of small sample biases, we also compute the population variance ratios implied by the three models (not reported). For the Mehra-Prescott economy, the variance ratio for equity returns is severely biased downward in small samples, but the bond return variance ratios remain close to the small sample values. The same is true for the Campbell-Shiller economy where equity returns in population are also slightly positively correlated. Hence, both economies fail to generate both in population and in small samples persistence in bond returns. The exception is the Moody Investor economy where in population variance ratios are well over 1.0 for bond returns. Moreover, equity returns show some weak mean reversion in population.

[18] Practitioners often view these relative yields as indications of fundamental value and use them in tactical asset allocation models. Although we do not focus on them, univariate regressions of both bond and stock returns on inflation and nominal interest rates typically yield insignificant coefficients.

regression. The link is not exact, since we use a coupon bond, whereas Campbell and Shiller use continuously compounded zero coupon rates. Recent research by Bekaert, Hodrick and Marshall (1997), among others, suggests that this empirical finding may constitute a serious challenge for any model of risk.

The slope coefficients implied by the models are reported in Table 8. It is remarkable how well the three models seem to capture the (weak) predictability in the data. Of the 12 coefficients displayed in the table, only one (equity on term spread in the Campbell-Shiller economy) has the wrong sign, and only one coefficient is not within two standard errors from the sample moment (the bond return on term spread regression in the Moody Investor economy).

One possibility is that the good performance is driven by small sample effects. That is, since all of these regressions feature rather persistent regressors, the coefficients will be biased in small samples [see Stambaugh (1986)]. Hence, if our theoretical economies generate persistent term spreads and dividend yields, that may be enough to obtain similar regression results as in the data, even though there is little true predictability in population. We checked this by deriving population regression coefficients through simulation. For the Mehra-Prescott economy we find that the population coefficients are uniformly smaller than the small sample regression coefficients, the largest being the bond return on the term spread, yielding a slope coefficient of 0.313 (versus 2.137 in the data). Not surprisingly, the bond return regression coefficients are essentially zero in the Campbell-Shiller economy as we know there is no time-variation in the bond premium in this model. The other slope coefficients are similarly small. Although the Moody Investor generates the highest positive regression slopes that seem most consistent with the data, the population slopes are small. In fact, the regression slope of excess equity on term spreads is even negative. Essentially, the Moody Investor economy has a channel to generate substantial time-variation in risk premiums, but overall the price of risk is very smooth. Given the observed state variables during our sample (which includes the Depression years, and some major recessions in the seventies and eighties), the effect on measured predictability is, however, rather substantial.[19]

5.3 Conditional Risk Premiums

To potentially gain more power, we also produce an alternative test of the performance of the various models with respect to predictability. Using the expressions for risk premiums on bonds and stocks in Equations (23) and (24), we can create unexpected returns predicted by the various models:

$$u_r^j(t+1) = R^j(t+1) - E[R^j(t+1)|I(t)] \qquad \text{with } j = s, b. \tag{52}$$

[19] Analogously, Bekaert, Hodrick and Marshall (2000) were also only able to explain the deviations from the Expectations Hypothesis by combining time-varying term premiums and small sample problems. As is the case here, term premiums in population were small and showed little variation.

If the model captures all relevant information about time-variation in expected returns, this unexpected return should be orthogonal to any pre-determined set of variables. We use as the instrument set a constant, the dividend yield, the term spread and also the nominal interest rate, since our closed-form solutions often predict a particular relation between risk premium and the nominal rate. In constructing the test, there are two sources of sampling error we have to take into account.

The first source arises from the small sample used in computing the moments themselves, and the second source is the uncertainty surrounding the true structural parameters. For a particular pre-estimated parameter configuration, we present a GMM-based predictability test taking both sources of standard errors into account. The predictability test is described in the Appendix.

This test is carried out in Table 9. The results are uniform across the three models. There is not enough power to reject the null that the model's unexpected equity returns are not predictable by the instruments, but for bond returns the null is rejected for all three models at the 1% level, with the test statistic value being lowest for the Moody Investor economy.

Table 9 also reports some characteristics, such as the minimum, maximum, mean, and volatility of the (gross) bond and equity return premiums implied by the model. There are no counterparts to these in the data. A naive approach to modeling expected returns would be to simply use the linear projections implied by the regression evidence. The last line reports some characteristics for the fitted values of a regression of returns onto our two yield instruments. However, the risk premiums obtained in this way seem excessively variable and often become negative. Generally, the model risk premiums behave more reasonably, in that their variation is more moderate and that equity premiums are always positive. As expected from the moment analysis above, bond premiums are always positive in the Campbell-Shiller and Moody Investor economy, but negative in the Mehra-Prescott model. It is here that the power of the Moody investor economy to generate time-varying prices of risk shows up most forcefully. Focussing on equity premiums, the sample variability in the other two models is negligible, but in the Campbell-Cochrane model it is 1.88%. Our regression-based procedure yields a variability of over 4%.

5.4 Excess Volatility Tests

Arguably the most powerful way to test for long-horizon predictability is to use the present value model directly. Intuitively, the price-dividend ratio should predict future dividend growth and future required rates of returns [Campbell and Shiller (1988) and Cochrane (1992)]. Its variability in the data (see Table 10) is estimated to be 7.70 with a standard error of 0.79. The challenge for our models is to match some salient features of bond and equity returns, whereas at the same time providing enough time-variation in discount rates to be able to match this large variability of price-dividend ratios.

Table 10 reports the implied variability of price-dividend ratios. Because of the close connection between discount rate and cash flow effects, it is not surprising to find that the Mehra-Prescott economy generates price-dividend variability that is much lower than in the other economies. However, the endogenous price-dividend variability generated by all three models remains starkly low. Since we failed to match the mean of the price-dividend ratio, and its variability will likely rise with the mean, we also report the coefficient of variation. In the data, the coefficient of variation equals 0.305. The models still fall considerably short of this, with the Campbell-Shiller economy, which has no equilibrium restrictions, being the one that comes the closest. Clearly, if we calibrate the models as we do, using dividend growth data and a close matching of the interest rate process, the excess variability puzzle of Shiller (1981), Kleidon (1986) and others remains.

These results generate somewhat of a puzzle. How can the Campbell-Shiller and Moody Investor economies generate realistic equity return variability, but fail to match price dividend ratio variability? After all, we have $R^s_{t+1} = \left(\frac{pd_{t+1}+1}{pd_t}\right)\exp(\pi_{t+1}+\Delta d_{t+1}) - 1$. Hence, if the price dividend ratio displays little variability, equity return variability is very close to the variability of nominal dividend growth rates. This is essentially what happens for the Mehra-Prescott economy. The equity return variability this model generates is not much higher than that of dividend growth. However, for the other economies, dividend variability accounts for less than 50% of the total variability of equity returns. The remainder is price dividend ratio variability, but the variability of equity returns is also increased by the positive correlation these models generate between dividend growth and price dividend ratios. In the Mehra-Prescott economy on the other hand, this variability is even less than zero for our sample. The actual numbers for the variance decomposition are reported in the final columns of Table 10. Intuitively, we would expect cash flows and prices to be positively correlated, but dividends may affect the discount rate process as well, and in an equilibrium context such as Lucas (1978), they may drive up discount rates, reducing prices. In our parameterization of the Mehra-Prescott economy, dividend growth is not directly priced (only X_t is priced) explaining the weak link between dividend growth and price dividend ratios.

6 Conclusion

In this paper, we have presented a stochastic valuation framework for pricing bonds and equities. We have first shown, in a tractable fashion, how the framework embeds a number of well-known pricing paradigms in both the term structure and equity pricing literature. In several examples we were able to derive closed-form solutions for equity and bond premiums. When confronted with the data, a three factor model can simultaneously match the equity premium and equity volatility, provided that either equilibrium restrictions are relaxed in a Campbell-Shiller like economy, or that a time-additive preference economy is generalized to an economy with preference shocks, as in the Moody Investor economy. The latter two models also generate upward sloping term structures on average, as is true in the data, but they still fail to match the variability of price-dividend ratios present in the data. Nevertheless, we took the data, in particular dividend growth, very seriously in our empirical exercise, despite their noisy nature. Lewellen and Shanken (2000) argue that parameter uncertainty can spuriously induce return predictability and price volatility even though the true process about which investors learn does not exhibit such predictability or volatility.

The basic model is flexible enough to be extended in many directions. First, our model has been extended to include a more generalized "external habit stock" in the style of Abel (1999).[20] Abel (1999) specifies an alternative and more general model of "external habit," in which the benchmark level of consumption can depend both on current and past consumption. His model embeds both the original "catching up with the Joneses" specification of Abel (1990) and the consumption externalities preferences of Gali (1994). Whereas Abel derives closed-form solutions for asset prices (bonds and stocks) under the assumption of i.i.d. consumption growth, his setup fits within our general model and we can accommodate more general dynamics for the state variables.[21]

Second, our model has been extended in several directions to explore the effects of dividend uncertainty on equity prices and examine the role it plays in accounting for endogenous asymmetric volatility in asset returns (the tendency of market volatility to rise more after bad news than after good news).[22]

Third, Brennan (1997) discusses how the evidence on predictability clashes with the practice of using a static CAPM for capital budgeting. In order to generate a time-varying discount rate, Brennan uses an empirical approach to first estimate the joint process for short and long-term interest rates, the market dividend yield, and the return on the market portfolio. He then performs a Monte Carlo simulation to estimate the expected return (and discount rate) on the market portfolio over a T-year horizon. The approach followed in this paper [and other related papers such as Ang and Liu (1999)] allows one to create discount functions that are consistent with predictability, change with the state of the economy, and use the information present in the term structure. That is, the present model allows one to construct an internally consistent model of time-varying risk premiums that follows directly from a simple, underlying theory. Such an approach could prove quite useful in capital budgeting applications.

Our approach has some disadvantages that provide substantial challenges for future work. First, we do not fully specify the general equilibrium that can support the kernel process, particularly on the monetary side. There are many ways to introduce

[20] This extension is available upon request from the authors.

[21] However, part of Abel's results do not assume log-normality, while ours do.

[22] This extension is available upon request from the authors. See Abel (1988), Campbell and Hentschel (1992) and Wu (2000).

money in a general equilibrium economy, but outside of putting real money balances in the utility function, it is difficult to retain tractability. Second, the preference structures allowed by our framework are not entirely general. An important class of models that does not fit into our framework are models with Kreps-Porteus preferences (1978) that allow the separation of risk aversion for timeless gambles from temporal elasticity of substitution. Campbell (1993) and Restoy and Weil (1998) have recently delivered tractable solutions for risk premiums in such models, relying on a log-linear approximation. Third, the permissible state variable dynamics are restrictive and do not allow for non-linearities except through stochastic volatility of the square root form. GARCH-type processes as in Bekaert (1996) or regime-switching processes as in Hung (1994) and Cecchetti, Lam and Mark (1990) cannot be accommodated. Such processes may be necessary to match the higher order moments of higher frequency return data.

References

[1] Abel, A., 1988, Stock prices under time-varying dividend risk: An exact solution in an infinite-horizon general equilibrium model, Journal of Monetary Economics 22, 375-393.

[2] Abel, A., 1990, Asset prices under habit formation and catching up with the Joneses, American Economic Review 80, 38-42.

[3] Abel, A. B., 1999, Risk premia and term premia in general equilibrium, Journal of Monetary Economics 43, 3-33.

[4] Ang, A., and J. Liu, 1999, A generalized earnings model of stock valuation, Working Paper, Stanford University.

[5] Ang, A., and M. Piazzesi, 2000, A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables, Working Paper, Columbia Business School.

[6] Bakshi, G. S., and Z. Chen, 1997, An alternative valuation model for contingent claims, Journal of Financial Economics 44, 123-165.

[7] Bakshi, G. S., and Z. Chen, 1998, Stock valuation in dynamic economies, Working Paper, University of Maryland.

[8] Bekaert, G., 1996, The time-variation of risk and return in foreign exchange markets: A general equilibrium perspective, Review of Financial Studies 9, 427-470.

[9] Bekaert, G., R. J. Hodrick and D. A. Marshall, 1997, The implications of first order risk aversion for asset market risk premiums, Journal of Monetary Economics 40, 3-39.

[10] Bekaert, G., R. J. Hodrick and D. A. Marshall, 2000, Peso problem explanations for term structure anomalies, Working Paper, Columbia Business School.

[11] Berk, J., R. Green and V. Naik, 1999, Optimal investment, growth options and security returns, Journal of Finance 54, 1153-1607.

[12] Bollerslev, T., and R. J. Hodrick, 1996, Financial market efficiency tests, in Pesaran, M. H. and M. R. Wickens, eds., The Handbook of Applied Econometrics, Vol. 1, Macroeconomics, Basil Blackwell, 415-458.

[13] Brennan, M. J., 1997, The term structure of discount rates, Financial Management 26, 81-90.

[14] Burnside, C. 1994, Hansen-Jagannathan bounds as classical tests of asset-pricing models, Journal of Business and Economic Statistics 12, 57-79.

[15] Campbell, J. Y., 1993, Intertemporal asset pricing without consumption data, American Economic Review 83, 487-512.

[16] Campbell, J. Y., and Cochrane, J. H., 1999, By force of habit: A consumption based explanation of aggregate stock market behavior, Journal of Political Economy 107, 205-251.

[17] Campbell, J. Y. and L. Hentschel, 1992, No news is good news: An asymmetric model of changing volatility in stock returns, Journal of Financial Economics 31, 281-318.

[18] Campbell, J. Y., and R. J. Shiller, 1988, The dividend price ratio and expectations of future dividends and discount factors, Review of Financial Studies 1, 195-228.

[19] Campbell, J. Y., and R. J. Shiller, 1991, Yield spreads and interest rate movements: A bird's eye view, Review of Economic Studies 58, 495-514.

[20] Cecchetti S., P.S. Lam and N. Mark, 1990, Mean reversion in equilibrium asset prices, American Economic Review 80, 398-418.

[21] Cochrane, J. H., 1992, Explaining the variance of price-dividend ratios, The Review of Financial Studies 5, 243-280.

[22] Cole, K., J. Helwege and D. Laster, 1996, Stock market valuation indicators: Is this time different?, Financial Analysts Journal 52, 56-64.

[23] Constantinides, G. M., 1992, A theory of the nominal term structure of interest rates, Review of Financial Studies 5, 531-552.

[24] Cox, J. C., J. E. Ingersoll and S. A. Ross, 1985, A theory of the term structure of interest rates, Econometrica 53, 385-408.

[25] Dai, Q., and K. J. Singleton, 2000, Specification analysis of affine term structure models, Journal of Finance, Forthcoming.

[26] Duffie, D., and R. Kan, 1996, A yield-factor model of interest rates, Mathematical Finance 6, 379-406.

[27] Fama, E., and K. French, 1989, Business conditions and expected returns on stocks and bonds, Journal of Financial Economics 25, 23-49.

[28] Gali, J., 1994, Keeping up with the Joneses: Consumption externalities, portfolio choice, and asset prices, Journal of Money, Credit, and Banking 26, 1-8.

[29] Hagiwara, M., and M. A. Herce, 1997, Risk aversion and stock price sensitivity to dividends, American Economic Review 87, 738-745.

[30] Hansen, L. P., 1982, Large sample properties of generalized method of moments estimators, Econometrica 50, 1029-1054.

[31] Hansen, L. P., and R. Jagannathan, 1991, Implications of security market data for models of dynamic economies, Journal of Political Economy 99, 225-262.

[32] Harrison, J. M., and D. Kreps, 1979, Martingales and arbitrage in multi-period securities markets, Journal of Economic Theory 20, 381-408.

[33] Harvey, C. R., 1989, Time-varying conditional covariances in tests of asset pricing models, Journal of Financial Economics 24, 289-317.

[34] Ho, T., and S. Lee, 1986, Term structure movements and pricing interest rate contingent claims, Journal of Finance 41, 1011-1029.

[35] Hung, M., 1994, The interaction between nonexpected utility and asymmetric market fundamentals, Journal of Finance 49, 325-343.

[36] Keim, D. B., and R. F. Stambaugh, 1986, Predicting returns in bond and stock markets, Journal of Financial Economics 17, 357-390.

[37] Kleidon, A. W., 1986, Variance bounds tests and stock price valuation models, Journal of Political Economy 94, 953-1001.

[38] Kocherlakota, 1996, The equity premium: It's still a puzzle, Journal of Economic Literature 304, 42-71.

[39] Kreps, D., and E. Porteus, 1978, Temporal resolution of uncertainty and dynamic choice theory, Econometrica 46, 185-200.

[40] Labadie, P., 1989, Stochastic inflation and the equity premium, Journal of Monetary Economics 24, 277-298.

[41] Lewellen, J., and J. Shanken, 2000, Estimation risk, market efficiency, and the predictability of stock returns, Working Paper, MIT.

[42] Lucas, R. E., Jr., 1978, Asset prices in an exchange economy, Econometrica 46, 1426-1446.

[43] Mehra, R., and E. C. Prescott, 1985, The equity premium: A puzzle, Journal of Monetary Economics 15, 145-161.

[44] Pearson, N., and T. Sun, 1994, An empirical examination of the Cox, Ingersoll, and Ross model of the term structure of interest rates using the method of maximum likelihood, Journal of Finance 54, 929-959.

[45] Poterba, J., and L. Summers, 1988, Mean reversion in stock prices: Evidence and implications, Journal of Financial Economics 22, 27-59.

[46] Restoy, F., and P. Weil, 1998, Approximate equilibrium asset prices, NBER Working Paper 6611.

[47] Richard, S., 1978, An arbitrage model of the term structure of interest rates, Journal of Financial Economics 6, 33-57.

[48] Sharpe, W. F., 1990, Investor wealth measures and expected return, in Quantifying the market risk premium phenomenon for investment decision making, The Institute of Chartered Financial Analysts, 29-37.

[49] Shiller, R. J., 1981, Do stock prices move too much to be justified by subsequent changes in dividends? American Economic Review 71, 421-426.

[50] Stambaugh, R., 1986, Bias in regression with lagged stochastic regressions, Working paper.

[51] Vasicek, O., 1977, An equilibrium characterization of the term structure, Journal of Financial Economics 5, 177-188.

[52] Weil, P., 1989, The equity premium puzzle and the risk free rate puzzle, Journal of Monetary Economics 24, 401-421.

[53] Wu, G., 2000, The determinants of asymmetric volatility, Review of Financial Studies, Forthcoming.

Appendix

In this Appendix we derive the general solutions for the pricing of bonds and equities presented in Propositions 1 and 2. We begin by defining the Hadamard Product, denoted by $\odot$. The use of this operator is solely for ease of notational complexity.

Definition: Suppose $A = (a_{ij})$ and $B = (b_{ij})$ are each $N \times N$ matrices. Then $A \odot B = C$, where $C = (c_{ij}) = (a_{ij} b_{ij})$ is an $N \times N$ matrix. Similarly, suppose $a = (a_i)$ is an N-dimensional column vector and $B = (b_{ij})$ is an $N \times N$ matrix. Then $a \odot B = C$, where $C = (c_{ij}) = (a_i b_{ij})$ is an $N \times N$ matrix. Again, suppose $a = (a_j)$ is an N-dimensional row vector and $B = (b_{ij})$ is an $N \times N$ matrix. Then $a \odot B = C$, where $C = (c_{ij}) = (a_j b_{ij})$ is an $N \times N$ matrix. Finally, suppose $a = (a_i)$, and $b = (b_i)$ are $N \times 1$ vectors. Then $a \odot b = C$, where $C = (c_i) = (a_i b_i)$ is an $N \times 1$ vector.

We shall find it useful to prove several lemmas. Let c be an $N \times 1$ vector and let α be a scalar.

Lemma 1: $Var_t(c^{'} Y_{t+1}) = \left[(\Sigma_F^{'} c) \odot (\Sigma_F^{'} c) \right]^{'} Y_t + c^{\prime} \Sigma_H \Sigma_H^{'} c.$
Proof:

$$\begin{aligned}
Var_t(c^{'} Y_{t+1}) &= c^{\prime} \left[(\Sigma_F F_t + \Sigma_H) (\Sigma_F F_t + \Sigma_H)^{'} \right] c \\
&= c^{\prime} \left[\Sigma_F F_t F_t^{'} \Sigma_F^{'} + \Sigma_F F_t \Sigma_H^{'} + \Sigma_H F_t^{'} \Sigma_F^{'} + \Sigma_H \Sigma_H^{'} \right] c \\
&= c^{\prime} \left[\Sigma_F (Y_t \odot I) \Sigma_F^{'} + \Sigma_H \Sigma_H^{'} \right] c \\
&= \left[\left(\Sigma_F^{'} c \right) \odot \left(\Sigma_F^{'} c \right) \right]^{'} Y_t + c^{\prime} \Sigma_H \Sigma_H^{'} c
\end{aligned}$$

where we use the conditions in (6) and the properties of the $\odot$ operator to simplify the expression.

Lemma 2: $Var_t(m_{t+1}) = (\Sigma_{mf} \odot \Sigma_{mf})^{'} Y_t + \Sigma_m^{'} \Sigma_m.$
Proof:

$$\begin{aligned}
Var_t(m_{t+1}) &= \left(\Sigma_{mf}^{'} F_t + \Sigma_m^{'} \right) \left(\Sigma_{mf}^{'} F_t + \Sigma_m^{'} \right)^{'} \\
&= \Sigma_{mf}^{'} F_t F_t^{'} \Sigma_{mf} + \Sigma_{mf}^{'} F_t \Sigma_m + \Sigma_m^{'} F_t^{'} \Sigma_{mf} + \Sigma_m^{'} \Sigma_m \\
&= \Sigma_{mf}^{'} (Y_t \odot I) \Sigma_{mf} + \Sigma_m^{'} \Sigma_m \\
&= (\Sigma_{mf} \odot \Sigma_{mf})^{'} Y_t + \Sigma_m^{'} \Sigma_m
\end{aligned}$$

where we use the conditions in (6) and the properties of the $\odot$ operator to simplify the expression.

Lemma 3: $Cov_t(c^{'} Y_{t+1}, m_{t+1}) = c^{'} \left[\left(\Sigma_{mf}^{'} \odot \Sigma_F \right) Y_t + \Sigma_H \Sigma_m \right].$

Proof:

$$
\begin{aligned}
Cov_t(c^{'}Y_{t+1}, m_{t+1}) &= c^{'}\left[(\Sigma_F F_t + \Sigma_H)\left(\Sigma_{mf}^{'}F_t + \Sigma_m^{'}\right)^{'}\right] \\
&= c^{'}\left[\Sigma_F F_t F_t^{'}\Sigma_{mf} + \Sigma_F F_t \Sigma_m + \Sigma_H F_t^{'}\Sigma_{mf} + \Sigma_H\Sigma_m\right] \\
&= c^{'}\left[\Sigma_F\left(Y_t \odot I\right)\Sigma_{mf} + \Sigma_H\Sigma_m\right] \\
&= c^{'}\left[\left(\Sigma_{mf}^{'} \odot \Sigma_F\right)Y_t + \Sigma_H\Sigma_m\right].
\end{aligned}
$$

where we use the conditions in (6) and the properties of the $\odot$ operator to simplify the expression.

Lemma 4: $E_t\left[\exp\left(\alpha + c^{'}Y_{t+1} + m_{t+1}\right)\right] = \exp\left(g_0 + g'Y\right),$
where:

$$
\begin{aligned}
g_0 &= \alpha + \mu_m + \frac{1}{2}\Sigma_m^{'}\Sigma_m + c^{'}\left(\mu + \Sigma_H\Sigma_m\right) + \frac{1}{2}c'\Sigma_H\Sigma_H^{'}c \\
g' &= \Gamma_m^{'} + \frac{1}{2}\left(\Sigma_{mf} \odot \Sigma_{mf}\right)^{'} + c^{'}\left[A + \left(\Sigma_{mf}^{'} \odot \Sigma_F\right)\right] + \frac{1}{2}\left(\Sigma_F^{'}c \odot \Sigma_F^{'}c\right)^{'}.
\end{aligned}
$$

Proof: By log-normality,

$$
E_t\left[\exp\left(\alpha + c^{'}Y_{t+1} + m_{t+1}\right)\right] = \exp\left[E_t\left(\alpha + c^{'}Y_{t+1} + m_{t+1}\right) + \frac{1}{2}Var_t\left(c^{'}Y_{t+1} + m_{t+1}\right)\right].
$$

We can first write:

$$
E_t\left(\alpha + c^{'}Y_{t+1} + m_{t+1}\right) = \alpha + c^{'}\left(\mu + AY_t\right) + \mu_m + \Gamma_m^{'}Y_t.
$$

In addition,

$$
\begin{aligned}
Var_t\left(c^{'}Y_{t+1} + m_{t+1}\right) &= Var_t\left(c^{'}Y_{t+1}\right) + Var_t\left(m_{t+1}\right) + 2Cov_t\left(c^{'}Y_{t+1}, m_{t+1}\right) \\
&= \left[\left(\Sigma_F^{'}c\right) \odot \left(\Sigma_F^{'}c\right)\right]^{'}Y_t + c'\Sigma_H\Sigma_H^{'}c + \left(\Sigma_{mf} \odot \Sigma_{mf}\right)^{'}Y_t + \Sigma_m^{'}\Sigma_m \\
&\quad +2c^{'}\left[\left(\Sigma_{mf}^{'} \odot \Sigma_F\right)Y_t + \Sigma_H\Sigma_m\right] \\
&= c'\Sigma_H\Sigma_H^{'}c + \Sigma_m^{'}\Sigma_m + 2c^{'}\Sigma_H\Sigma_m \\
&\quad + \left[\left(\Sigma_{mf} \odot \Sigma_{mf}\right)^{'} + 2c^{'}\left(\Sigma_{mf}^{'} \odot \Sigma_F\right) + \left(\Sigma_F^{'}c \odot \Sigma_F^{'}c\right)^{'}\right]Y_t,
\end{aligned}
$$

where we apply lemmas 1, 2, and 3, and the properties of the $\odot$ operator.

Thus,

$$
E_t\left(c^{'}Y_{t+1} + m_{t+1}\right) + \frac{1}{2}Var_t\left(c^{'}Y_{t+1} + m_{t+1}\right) = g_0 + g^{'}Y_t.
$$

Proof of Proposition 1

The derivation begins by guessing that the solution for the log of bond prices equals:

$$p_{n,t} = a_n + A_n^{'} Y_t. \tag{A.1}$$

We shall verify that the guess is indeed correct.

Under the nominal pricing kernel, the time t value of an n-year bond must satisfy:

$$\begin{aligned} \exp(p_{n,t}) &= E_t\left[\exp(\hat{m}_{t+1} + p_{n-1,t+1})\right] \\ &= E_t\left[\exp(m_{t+1} - e_\pi^{'} Y_{t+1} + p_{n-1,t+1})\right] \end{aligned} \tag{A.2}$$

where we write the nominal kernel as the real kernel minus inflation.

Using our "guess" for the form of $p_{n,t}$, we can then write:

$$\exp(p_{n,t}) = E_t\left[\exp\left(a_{n-1} + \left(A_{n-1}^{'} - e_\pi^{'}\right) Y_{t+1} + m_{t+1}\right)\right]. \tag{A.3}$$

Using lemma 4, with $\alpha = a_{n-1}$ and $c = (A_{n-1} - e_\pi)$, we have:

$$\exp(p_{n,t}) = \exp\left(g_0 + g'Y\right), \tag{A.4}$$

with:

$$\begin{aligned} g_0 &= a_{n-1} + \mu_m + \frac{1}{2}\Sigma_m^{'}\Sigma_m + \left(A_{n-1} - e_\pi\right)^{'}\left[\mu + \Sigma_H \Sigma_m\right] \\ &\quad + \frac{1}{2}\left(A_{n-1} - e_\pi\right)^{'} \Sigma_H \Sigma_H^{'} \left(A_{n-1} - e_\pi\right), \\ g^{'} &= \Gamma_m^{'} + \frac{1}{2}\left(\Sigma_{mf} \odot \Sigma_{mf}\right)^{'} + \left(A_{n-1} - e_\pi\right)^{'}\left[A + \Sigma_{mf}^{'} \odot \Sigma_F\right] \\ &\quad + \frac{1}{2}\left[\Sigma_F^{'}\left(A_{n-1} - e_\pi\right) \odot \Sigma_F^{'}\left(A_{n-1} - e_\pi\right)\right]^{'}. \end{aligned} \tag{A.5}$$

Thus, $p_{n,t} = g_0 + g^{'} Y$, and our guess is verified by setting $a_n = g_0$, and $A_n^{'} = g^{'}$.

Proof of Proposition 2

From equation. (12), the price-dividend ratio, pd_t, can be expressed as:

$$pd_t = \frac{V_t}{D_t} = E_t\left\{\sum_{n=1}^{\infty} \exp\left[\sum_{j=1}^{n}\left(m_{t+j} + \Delta d_{t+j}\right)\right]\right\}. \tag{A.6}$$

Define $q_{n,t} = E_t\left\{\exp\left[\sum_{j=1}^{n}\left(m_{t+j} + \Delta d_{t+j}\right)\right]\right\} = E_t\left\{\exp\left[\sum_{j=1}^{n}\left(m_{t+j} + e_d^{'} Y_{t+j}\right)\right]\right\}$, for $n = 1, 2, \ldots$. Thus,

$$pd_t = \sum_{n=1}^{\infty} q_{n,t}. \tag{A.7}$$

We will now prove by mathematical induction that $q_{n,t}$ can be written as:

$$q_{n,t} = \exp\left(b_n + B_n^{'} Y_t\right), \tag{A.8}$$

where b_n and B_n are defined by the difference equations in (14).

First, we show that $q_{1,t} = \exp(m_{t+1} + e_d^{'} Y_{t+1})$ can be written in this affine form as $q_{1,t} = \exp\left(b_1 + B_1^{'} Y_t\right)$. This is clear, since we can use lemma 4 by setting $\alpha = 0$ and $c = e_d$. The proposed solution holds provided:

$$\begin{aligned} b_1 &= \mu_m + \frac{1}{2}\Sigma_m^{'}\Sigma_m + e_d^{'}\left(\mu + \Sigma_H \Sigma_m\right) + \frac{1}{2} e_d^{'} \Sigma_H \Sigma_H^{'} e_d \\ B_1^{'} &= \Gamma_m^{'} + \frac{1}{2}\left(\Sigma_{mf} \odot \Sigma_{mf}\right)^{'} + e_d^{'}\left[A + \left(\Sigma_{mf}^{'} \odot \Sigma_F\right)\right] + \frac{1}{2}\left(\Sigma_F^{'} e_d \odot \Sigma_F^{'} e_d\right)^{'}. \end{aligned} \tag{A.9}$$

Thus, we have verified our solution for the case of $n = 1$.

Now, assume that $q_{n-1,t} = \exp\left(b_{n-1} + B_{n-1}^{'} Y_t\right)$. We now show that $q_{n,t} = \exp\left(b_n + B_n^{'} Y_t\right)$.

$$\begin{aligned} q_{n,t} &= E_t\left\{\exp\left[\sum_{j=1}^{n}\left(m_{t+j} + e_d^{'} Y_{t+j}\right)\right]\right\} \\ &= E_t\left\{\exp\left[\left(m_{t+1} + e_d^{'} Y_{t+1}\right) + \sum_{j=1}^{n-1}\left(m_{t+1+j} + e_d^{'} Y_{t+1+j}\right)\right]\right\} \\ &= E_t\left\{E_{t+1}\left[\exp\left(m_{t+1} + e_d^{'} Y_{t+1}\right) \cdot \exp\left(\sum_{j=1}^{n-1}\left(m_{t+1+j} + e_d^{'} Y_{t+1+j}\right)\right)\right]\right\} \\ &= E_t\left\{\exp\left(m_{t+1} + e_d^{'} Y_{t+1}\right) \cdot E_{t+1}\left[\exp\left(\sum_{j=1}^{n-1}\left(m_{t+1+j} + e_d^{'} Y_{t+1+j}\right)\right)\right]\right\} \\ &= E_t\left\{\exp\left(m_{t+1} + e_d^{'} Y_{t+1} + q_{n-1,t+1}\right)\right\} \\ &= E_t\left\{\exp\left(b_{n-1} + \left(B_{n-1} + e_d\right)^{'} Y_{t+1} + m_{t+1}\right)\right\}. \end{aligned} \tag{A.10}$$

We can use lemma 4 by setting $\alpha = b_{n-1}$ and $c = (B_{n-1} + e_d)$. The proposed solution holds provided:

$$\begin{aligned} b_n &= b_{n-1} + \mu_m + \frac{1}{2}\Sigma_m^{'}\Sigma_m + \left(B_{n-1} + e_d\right)^{'}\left[\mu + \Sigma_H \Sigma_m\right] \\ &\quad + \frac{1}{2}\left(B_{n-1} + e_d\right)^{'} \Sigma_H \Sigma_H^{'} \left(B_{n-1} + e_d\right), \\ B_n^{'} &= \Gamma_m^{'} + \frac{1}{2}\left(\Sigma_{mf} \odot \Sigma_{mf}\right)^{'} + \left(B_{n-1} + e_d\right)^{'}\left[A + \Sigma_{mf}^{'} \odot \Sigma_F\right] \\ &\quad + \frac{1}{2}\left[\Sigma_F^{'}\left(B_{n-1} + e_d\right) \odot \Sigma_F^{'}\left(B_{n-1} + e_d\right)\right]^{'}. \end{aligned} \tag{A.11}$$

We have therefore verified the solution in Proposition 2.

Demonstrating That the Three-Factor Model Falls Within the General Affine Class

The model outlined in system (16) is a special case of the general affine class where the following parametric definitions are applied:
$I = 1$ (homoskedastic model)

$$\mu = \begin{pmatrix} \mu_d \\ \mu_x \\ \mu_\pi \end{pmatrix} \quad A = \begin{pmatrix} \rho_d & g_x & 0 \\ g_d & \rho_x & 0 \\ 0 & 0 & \rho_\pi \end{pmatrix} \quad \Gamma_m = \begin{pmatrix} \lambda_d \\ \lambda_x \\ 0 \end{pmatrix}$$

$$\Sigma_F = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sigma_\pi \end{pmatrix} \quad \Sigma_H = \begin{pmatrix} \sigma_d & \sigma_{dx} & 0 \\ 0 & \sigma_x & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \Sigma_{mf} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \Sigma_m = \begin{pmatrix} s_d \\ s_x \\ 0 \end{pmatrix}.$$

$I = 0$ (heteroskedastic model)

$$\mu = \begin{pmatrix} \mu_d \\ \mu_x \\ \mu_\pi \end{pmatrix} \quad A = \begin{pmatrix} \rho_d & g_x & 0 \\ g_d & \rho_x & 0 \\ 0 & 0 & \rho_\pi \end{pmatrix} \quad \Gamma_m = \begin{pmatrix} \lambda_d \\ \lambda_x \\ 0 \end{pmatrix}$$

$$\Sigma_F = \begin{pmatrix} 0 & v_{dx} & 0 \\ 0 & v_x & 0 \\ 0 & 0 & \sigma_\pi \end{pmatrix} \quad \Sigma_H = \begin{pmatrix} \sigma_d & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \Sigma_{mf} = \begin{pmatrix} 0 \\ v_{mx} \\ 0 \end{pmatrix} \quad \Sigma_m = \begin{pmatrix} s_d \\ 0 \\ 0 \end{pmatrix}.$$

Description of the GMM-Based Predictability Test

The predictability test can be described as follows. Denote the orthogonality conditions used to estimate Ψ as $g_{1T}(\Psi)$ and the orthogonality conditions we wish to test as $g_{2T}(\Psi)$. By the Mean Value Theorem,

$$g_{2T}(\hat{\Psi}) \overset{a.s.}{=} g_{2T}(\Psi_0) + D_{2T}(\Psi_0) \cdot \left(\hat{\Psi} - \Psi_0\right), \tag{53}$$

where Ψ_0 is the true parameter vector, and $\hat{\Psi}$ is the estimated parameter vector, and

$$D_{2T}(\Psi_0) = \frac{\partial g_{2T}(\Psi_0)}{\partial \Psi}. \tag{54}$$

Since we estimate $\hat{\Psi}$ from the first set of orthogonality conditions:

$$\hat{\Psi} - \Psi_0 \overset{a.s.}{=} -\left(A_{11} D_{1T}\right)^{-1} A_{11} \cdot g_{1T}(\Psi_0), \tag{55}$$

with

$$\begin{aligned} D_{1T} &= \frac{\partial g_{1T}(\Psi_0)}{\partial \Psi}, \\ A_{11} &= D_{1T}^{'} \cdot S_{11}^{-1}, \end{aligned} \tag{56}$$

where S_{11} is the spectral density at frequency zero of the orthogonality conditions g_{1T}. But then,

$$g_{2T}(\hat{\Psi}) \overset{a.s.}{=} M g_T(\Psi_0), \tag{57}$$

with

$$M = \left[-D_{2T} \cdot (A_{11} D_{1T})^{-1} A_{11}', \qquad I\right] \tag{58}$$

Since we can assume that $\sqrt{T} g_T(\Psi_0) \to N(0, S)$, where S is the spectral density at frequency zero of the orthogonality conditions, and

$$g_T(\Psi_0) = \left[g_{1T}(\Psi_0)', g_{2T}(\Psi_0)'\right]', \tag{59}$$

the statistic

$$T \cdot g_{2T}(\hat{\Psi})' \left[MSM'\right]^{-1} g_{2T}(\hat{\Psi}) \tag{60}$$

will have a $\chi^2(k)$ distribution under the null, where k is the number of moments considered in g_{2T}. In our case, $k = 4$, since we use four instruments and test bond and stock return predictability separately.

Table 1
Data Sources

Series	Symbol	Source	Availability
Nominal Stock Return	r^s_{t+1}	Ibbotson (S&P 500)	1926:96
Nominal Bond Return	r^b_{t+1}	Ibbotson (20 year bond)	1926:96
Nominal Interest rate	r_t	Ibbotson (one year T-bill)	1926:96
Inflation	π_t	Ibbotson	1926:96
Long Yield	lr_t	Board of Governors	1925:96
Real Dividend growth (end-of-period)	$\Delta d_{t,1}$	Own Computations	1927:96
Real Dividend Growth (additive)	$\Delta d_{t,2}$	Own Computations	1927:96
Price Dividend Ratio = 1/Dividend Yield	$pd_t = 1/dy_t$	Own Computations	1926:96
Term Spread	$lr_t - r_t$	Own Computations	1926:96

Note: Stock and bond returns, the nominal interest rate, inflation, the long yield, real dividend growth and the term spread are all in logs. To compute nominal dividend growth, assume c_t is the gross capital gain return over the year and i_t the income return ($i_t = D^n_{t+1}/Q_t$, with D^n_{t+1} the nominal dividend, and Q_t the price level). In the end-of-period case, the income return is computed assuming dividends are re-invested in the stock market. In the additive case, we simply add the dividends paid out during the year. Then, $D^n_{t+1}/D^n_t = (i_{t+1}/i_t)\, c_t$, and real dividend growth is $\Delta d_t = log(D^n_{t+1}/D^n_t) - \pi_t$.

Table 2
Empirical Properties of the Variables

State Variables			
	Dividend Growth (end-of-period)	Dividend Growth (additive)	Inflation
Mean	0.008 (0.016)	0.008 (0.016)	0.037 (0.005)
σ	0.137 (0.012)	0.123 (0.023)	0.031 (0.004)
rho	-0.098 (0.109)	0.185 (0.154)	0.922 (0.040)

Instruments			
	Dividend Yield	Term Spread	Interest Rate
Mean	0.044 (0.002)	0.009 (0.002)	0.040 (0.005)
σ	0.015 (0.002)	0.013 (0.001)	0.032 (0.004)
rho	0.667 (0.094)	0.735 (0.051)	0.906 (0.042)

Notes: All variables are in logs, except for the dividend yield. All moments were estimated using GMM [Hansen (1982)] allowing for one Newey-West lag.

Table 3

Example of Calibration of γ Parameter in the Mehra-Prescott Economy

γ	Mean Equity Premium
Estimate = 0.283	---
2	1.51
3	2.78
5.55	**6.09**
10	11.90
Data	6.14
(s.e.)	(2.40)

Notes: We estimate the parameter set for the Mehra-Prescott economy $\Psi = [\beta, \sigma_d, \mu_\pi, \rho_\pi, \sigma_\pi, \mu_x, \rho_x, \sigma_x]'$ using 9 moments of dividend growth, the nominal rate and inflation in a GMM-system. The estimated parameter values are used to infer the state variables from the data and to compute the mean equity premium (in logs). The last line reports the data moment with a GMM-based standard error between parentheses.

Table 4

Equity Characteristics of the Three Example Economies

	Mean Equity Premium	Mean pd_t	Equity Variability
MP-model (γ=5.55)	6.09	16.81	12.64
CS-model (s_d=-0.45)	6.31	16.34	19.36
MI-model (γ=2.60)	6.18	16.72	18.56
Data	6.14	25.23	19.58
(s.e.)	(2.40)	(1.20)	(2.16)

Notes: MP stands for Mehra-Prescott, CS for Campbell-Shiller and MI for Moody Investor. The parameter γ is chosen to roughly match the mean equity premium in the Mehra-Prescott Economy and in the Moody Investor economy. In the Campbell-Shiller economy, the parameter s_d is similarly calibrated. We report the mean equity excess return and its volatility and the mean price dividend ratio computed over the actual data sample, with the state variables inferred from data on dividend growth, nominal rates, and inflation. The last line reports the data moment with a GMM-based standard error between parentheses.

Table 5

Bond Characteristics of the Three Example Economies

	Mean Term Spread	Mean Bond Premium	Bond Variability	Bond/Equity Correlation
MP-model (γ=5.55)	-0.19	-0.57	7.38	-0.04
CS-model (s_d=-0.45)	0.35	0.44	7.80	0.34
MI-model (γ=2.60)	0.59	0.54	9.81	0.10
Data	0.95	0.90	7.82	0.189
(s.e.)	(0.21)	(0.92)	(0.77)	(0.089)

Notes: MP stands for Mehra-Prescott, CS for Campbell-Shiller and MI for Moody Investor. The parameter γ is chosen to roughly match the mean equity premium in the Mehra-Prescott Economy and in the Moody Investor economy. In the Campbell-Shiller economy, the parameter s_d is similarly calibrated. We report the term spread, the mean bond excess return and its volatility and the correlation between bond and equity returns computed over the actual data sample, with the state variables inferred from data on dividend growth, nominal rates, and inflation. The last line reports the data moment with a GMM-based standard error between parentheses.

Table 6

Skewness/Kurtosis for the Three Example Economies

	Equity		Bonds	
	Skewness	Kurtosis	Skewness	Kurtosis
MP-model (γ=5.55)	-1.19	3.96	0.73	1.56
CS-model (s_d =-0.45)	-0.69	3.22	0.54	1.05
MI-model (γ=2.60)	-1.13	6.43	0.84	1.59
Data	-0.906 (0.283)	1.187 (0.780)	1.157 (0.294)	1.716 (1.096)

Notes: MP stands for Mehra-Prescott, CS for Campbell-Shiller and MI for Moody Investor. The parameter γ is chosen to roughly match the mean equity premium in the Mehra-Prescott Economy and in the Moody Investor economy. In the Campbell-Shiller economy, the parameter s_d is similarly calibrated. We report skewness and *excess* kurtosis for equity and bond returns computed over the actual data sample, with the state variables inferred from data on dividend growth, nominal rates, and inflation. The last line reports the data moment with a GMM-based standard error between parentheses.

Table 7

Variance Ratios for the Three Example Economies

	Equity		Bonds	
	VR(5)	VR(10)	VR(5)	VR(10)
MP-model (γ=5.55)	0.92	0.79	0.76	0.75
CS-model (s_d =-0.45)	0.50	0.39	0.77	0.74
MI-model (γ=2.60)	0.52	0.43	0.96	1.03
Data (s.e.)	0.726 (0.137)	0.734 (0.150)	1.397 (0.260)	1.885 (0.416)

Notes: MP stands for Mehra-Prescott, CS for Campbell-Shiller and MI for Moody Investor. The parameter γ is chosen to roughly match the mean equity premium in the Mehra-Prescott Economy and in the Moody Investor economy. In the Campbell-Shiller economy, the parameter s_d is similarly calibrated. VR(k) stands for variance ratio computed using k autocorrelations of the underlying process. We report variance ratios using 5 or 10 autocorrelations for stock and bond returns computed over the actual data sample, with the state variables inferred from data on dividend growth, nominal rates, and inflation. The last line reports the corresponding variance ratios in the data with the standard errors computed by estimating the correlations jointly in a GMM framework. We use 11 Newey - West lags for this estimation.

Table 8

Predictability Properties for the Three Example Economies

	Equity		Bonds	
	Excess Dividend Yield	Term Spread	Excess Dividend Yield	Term Spread
MP-model (γ=5.55)	0.370	0.634	0.273	0.909
CS-model (s_d =-0.45)	0.294	0.028	0.350	1.427
MI-model (γ=2.60)	0.864	2.640	0.515	6.376
Data	0.875	2.936	0.179	2.137
(s.e.)	(0.595)	(1.698)	(0.276)	(0.622)

Notes: MP stands for Mehra-Prescott, CS for Campbell-Shiller and MI for Moody Investor. The parameter γ is chosen to roughly match the mean equity premium in the Mehra-Prescott Economy and in the Moody Investor economy. In the Campbell-Shiller economy, the parameter s_d is similarly calibrated. We report slope coefficients from univariate regressions of equity or bond excess returns onto excess dividend yield or term spreads. The regression variables are computed over the actual data sample, with the state variables inferred from data on dividend growth, nominal rates, and inflation. The last line reports the coefficients actually obtained in the data using heteroskedasticity-robust standard errors.

Table 9

Analytical Risk Premiums: Tests and Properties

	Bond Returns					Equity Returns				
	Test	Min.	Max.	Mean	Vol.	Test	Min.	Max.	Mean	Vol.
MP-model	14.53 (0.006)	-0.0039	0.0000	-0.0010	0.0009	5.70 (0.223)	0.0741	0.0861	0.0782	0.0028
CS-model	14.54 (0.0058)	0.0080	0.0106	0.0100	0.0005	5.66 (0.226)	0.0836	0.0966	0.0882	0.0029
MI-model	13.55 (0.009)	0.0027	0.0182	0.0099	0.0021	5.94 (0.204)	0.0222	0.1555	0.0858	0.0187
Data		-0.054	0.086	0.009	0.031		-0.044	0.139	0.061	0.042

Notes: The column labelled "Test" reports the value of the test statistic described in section 5.3 and in the Appendix, which is distributed $\chi^2(4)$. The p-value is indicated between parentheses. The columns min., max., mean and vol. report these sample characteristics for the (gross) bond or equity return premiums implied by the model. MP stands for Mehra-Prescott Economy, CS for Campbell-Shiller Economy and MI for Moody Investor Economy. The last line reports the same properties for fitted values from multivariate regressions of stock or bond returns on the two "yield" instruments.

Table 10

Variability in Price Dividend Ratios

	Standard deviation	Coefficient of variation	Variance of equity return accounted for by price-dividend ratio	Variance of equity return accounted for by dividend growth
MP-model	0.662	0.039	14.30	95.03
CS-model	1.697	0.104	36.80	40.36
MI-model	1.116	0.067	19.73	45.88
Data (s.e.)	7.670 (0.792)	0.305 (0.148)	126.12	40.32

Notes: MP stands for Mehra-Prescott, CS for Campbell-Shiller and MI for Moody Investor. The first column reports the standard deviation (volatility) of price-dividend ratios, the second its coefficient of variation both for the three models and the data. The number between parentheses is a GMM-based standard error. For the coefficient of variation, the standard error is computed using the delta method, since it is a function of the first two moments. We could also view it as a function of the mean and the volatility, in which case the standard error is reduced to 0.030. The last two columns report the results of a variance docomposition of the logarithmic equity return into two components: $\ln(\frac{1+pd_{t+1}}{pd_t})$ and nominal dividend growth. The numbers are in percentages.