

Stock and Bond Returns with Moody Investors

Geert Bekaert*

Columbia University and NBER

Eric Engstrom

University of Michigan

Steve Grenadier

Stanford University and NBER

This Draft: 02 July 2004

JEL Classifications

G12, G15, E44

Keyphrases

Empirical Asset Pricing, Stock-Bond Correlation, Macroeconomic Factors

Abstract:

We present a tractable, linear model for the simultaneous pricing of stock and bond returns that incorporates stochastic risk aversion. In this model, analytic solutions for endogenous stock and bond prices and returns are readily calculated. After estimating the parameters of the model by GMM, we investigate a series of classic puzzles of the empirical asset pricing literature. In particular, our model is shown to jointly accommodate the mean and volatility of equity and long term bond risk premia as well as salient features of the nominal short rate, the dividend yield, and the term spread. Also, the model matches the evidence for predictability of excess stock and bond returns. However, the stock-bond return correlation implied by the model is somewhat higher than in the data.

*Corresponding author: Columbia Business School, 802 Uris Hall, 3022 Broadway, New York New York, 10027; ph: (212)-854-9156; fx: (212)-662-8474; gb241@columbia.edu.

We thank Pietro Veronesi for useful comments.

1 Introduction

Campbell and Cochrane (1999) identify slow countercyclical risk premiums as the key to explaining a wide variety of dynamic asset pricing phenomena within the context of a consumption-based asset-pricing model. They generate such risk premiums by adding a slow moving external habit to the standard power utility framework. Essentially, as we clarify below, their model generates countercyclical risk aversion. This idea has surfaced elsewhere as well. Sharpe (1990) and practitioners such as Persaud (see, for instance, Kumar and Persaud 2002) developed models of time-varying risk appetites to make sense of dramatic stock market movements.

The first contribution of this article is to present a very tractable, linear model that incorporates stochastic risk aversion. Because of the model's tractability, it becomes particularly simple to address a wider set of empirical puzzles than those considered by Campbell and Cochrane. Campbell and Cochrane match salient features of equity returns, including the equity premium, excess return variability and the variability of the price dividend ratio. They do so in a model where the risk free rate is constant. Instead, we embed a fully stochastic term structure into our model, and investigate whether the model can fit salient features of bond and stock returns simultaneously. Such over-identification is important, because previous models that match equity return moments often do so by increasing the variability of marginal rates of substitution to the point that a satisfactory fit with bond market data and risk free rates is no longer possible. Using the General Method of Moments, we find that our model can successfully fit many features of bond and stock return data together with important properties of the fundamentals. Our economy generates a persistent interest rate process, with the same variability as observed in the data, realistic equity and bond premia, excess stock and bond returns as variable as in the data, and realistic dividend yield and term spread processes.

The NBER Working Paper version of Campbell and Cochrane also considered a specification with a stochastic interest rate. While that model matched some salient features of interest rate data, being a one-factor model, it necessarily could not provide a fully satisfactory fit of term structure data. Moreover, the one shock nature of the model imposes too strong of a link between bond and stock returns, an issue not examined in Campbell and Cochrane. The results in Campbell and Cochrane (1995) were nevertheless suggestive that a joint modeling of bond and stock returns within their or a related framework might prove successful.

Once we model bond and stock returns jointly, a series of classic empirical puzzles becomes testable. First, Shiller and Beltratti (1992) point out that present value models with a constant risk premium imply a negligible correlation between stock and bond returns in contrast to the moderate positive correlation in the data. We expand on the present value approach by allowing for an endogenously determined stochastic risk premium. Second, Fama and French (1989) and Keim and Stambaugh (1986) find common predictable components in bond and equity returns. After estimating the parameters of the model to match the salient features of bond and stock returns alluded to above, we test how well the model fares with respect to these puzzles. Our model generates a bond-stock return correlation that is somewhat too high relative to the data but it matches the predictability evidence.

Along the way, we face several measurement issues. First, to price stocks, we must use dividends as the relevant cash flow process, but the representative agent (stock holder) has utility over consumption. We take this measurement problem seriously by considering several possible stochastic relations between dividends and aggregate consumption. This aspect of our analysis is related to a number of recent articles which have stressed the presence of a moving average component in the dividend process, see, for instance Bansal and Lundblad (2002) and Bansal and Yaron (2001).

Second, we measure consumption using aggregate nondurables and services data. Campbell and Cochrane (1999) fail to match the low correlation between consumption growth and stock returns, which might reflect the fact that aggregate consumption is not representative of the consumption of stockholders (see Mankiw and Zeldes (1991) and Ait-Sahalia, Parker and Yogo (2003)). While we considered alternative consumption measures, our model indeed simultaneously fits the salient features of stock and bond returns, while keeping the correlation between returns and fundamentals low, even using nondurables and services as the consumption measure.

Third, to convert from model output to the data, we use inflation as a state variable, but ensure that inflation is neutral: that is the Fisher hypothesis holds in our economy. This is important in interpreting our empirical results on the joint properties of bond and stock returns. More realistic modeling of the inflation process is a prime candidate for resolving the remaining failures of the model.

Our model also fits into a long series of recent attempts to break the tight link between consumption growth and the pricing kernel that is the main reason for the failure of the standard consumption – based asset pricing models. Santos and Veronesi (2000) add the consumption/labor income ratio

as a second factor to the kernel, Wei (2003) adds leisure services to the pricing kernel and models human capital formation, Piazzesi, Schneider and Tuzel (2003) and Lustig and Van Nieuwerburgh (2003) model the housing market to increase the dimensionality of the pricing kernel.

The remainder of the article is organized as follows. Section 1 presents the model. Section 2 derives closed-form expressions for bond prices and equity returns. Section 3 outlines our estimation procedure whereas Section 4 analyzes the estimation results, and the implications of the model at the estimated parameters. Section 5 tests how the model fares with respect to the interaction of bond and stock returns. In the conclusions, we summarize the implications of our work for future research and we discuss some recent papers that have also considered the joint modeling of bond and stock returns, but were written after earlier versions of this article had been widely circulated.

2 The “Moody” Investor Economy

2.1 Preferences

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

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

where C_t is aggregate consumption and H_t is an exogenous “external habit stock” with $C_t > H_t$.

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.² 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 local 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 ratio³. As the surplus ratio goes to zero, the consumer’s risk aversion goes to infinity. In our model, we view the inverse of the

²For empirical analyses of habit formation models, where habit depends on past consumption, see Heaton (1995) and Bekaert (1996).

³Of course, this not risk aversion defined over wealth which depends on the value function. The Appendix to Campbell and Cochrane (1995) examines the relation between “local” curvature and actual risk aversion, which depends on the sensitivity of consumption to wealth. In their model, actual risk aversion is simply a scalar multiple of local curvature. In the present article, we only refer to the local curvature concept, and slightly abuse terminology in calling it “risk aversion.”

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 $Q_t > 1$. As Q_t changes over time, the representative consumer / investor's moodiness changes.

The marginal rate of substitution in this model determines the real pricing kernel, which we denote by M_t . Taking the ratio of marginal utilities of time $t + 1$ and t , we obtain:

$$\begin{aligned} M_{t+1} &= \beta \frac{(C_{t+1}/C_t)^{-\gamma}}{(Q_{t+1}/Q_t)^{-\gamma}} \\ &= \beta \exp[-\gamma \Delta c_{t+1} + \gamma (q_{t+1} - q_t)], \end{aligned} \quad (2)$$

where $q_t = \ln(Q_t)$ and $\Delta c_t = \ln(C_t) - \ln(C_{t-1})$.

This model may better explain the predictability evidence than the standard model with power utility. 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:

$$\frac{\sigma_t(M_{t+1})}{E_t(M_{t+1})} = \sqrt{\exp[Var_t(m_{t+1}) - 1]}. \quad (3)$$

where $m_t = \ln(M_t)$. 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. Notice that the conditional variance of the pricing kernel is proportional to the sum of the conditional variance of consumption growth and the conditional variance of Q_t .

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 the affine asset pricing framework. The unobserved process for $q_t \equiv \ln(Q_t)$ is included as an element of the state vector. Although the intertemporal marginal rate of substitution determines the form of the real pricing kernel through Equation (2), we still have a choice on how to model Δc_t and q_t .

Since $Q_t > 1$, we model q_t according to the specification,

$$q_{t+1} = \mu_q + \rho_q q_t + \sigma_q \sqrt{q_t} \varepsilon_{t+1}^q, \quad (4)$$

where μ_q , ρ_q and σ_q are parameters and σ_q is parameterized as $\frac{1}{f} \sqrt{\mu_q (1 - \rho_q^2)}$. It is easily shown that f is the ratio of the unconditional mean to the unconditional standard deviation of q_t . By bounding f below at unity, we ensure that q_t is ‘usually’ positive (under our subsequent estimates, q_t is positive in more than 95% of simulated draws). The fact that we model q_t as a square root process makes the conditional variance of the pricing kernel depend positively on the level of Q_t . However, to determine its full impact we must also model consumption growth.

2.2 Fundamentals Processes

When taking a Lucas-type economy to the data, the identity of the representative agent and the representation of the endowment process become critical. Because we price equities in this article, dividend growth is a state variable. Whereas in the simple Lucas model $C_t = D_t$, this cannot be true in the data. Consumption is financed by many sources of income (especially labor income) not represented in aggregate dividends. Moreover, many agents in the economy do not hold stocks at all, and their consumption may not be relevant for equity pricing. Campbell (1993) used the uncertainty about how to measure consumption as motivation to substitute consumption out of the budget constraint.

To deal with this measurement problem, we proceed in several steps. First, we introduce multiple dividend processes that break the link between C_t and D_t in the Lucas model. Eventually, this leads to a representation of consumption as dividends times the consumption dividend ratio. The main empirical issue then becomes whether the consumption dividend ratio is stationary or not. A number of articles have modeled the joint dynamics of dividends and consumption, implicitly assuming two separate unit roots. Bansal, Dittmar, and Lundblad (2002) recently argue that dividends and consumption are cointegrated, but with a cointegrating vector that differs from $[1, -1]$. Second, whereas our base specification works with aggregate nondurables and services consumption as the consumption measure, we checked the robustness of the model to an alternative measure of consumption that attempts to approximate the consumption of the stockholder. In particular, we let the stockholder consumption be a weighted average of luxury consumption and ‘other’ consumption

with the weighting equal to the stock market participation rate based on Ameriks and Zeldes (2001). However, our model does not perform noticeably better with this consumption measure and we do not report these results to conserve space. Section 2.2.1 details the modeling of consumption and dividend growth. Finally, to link a real consumption model to the nominal data, we must make assumptions about the inflation process, which we describe in Section 2.2.2.

2.2.1 Consumption and Dividends

In the original Lucas (1978) model, a dividend – producing ‘tree’ finances all consumption. To break the link between dividends on the aggregate stock market and aggregate consumption (even when limited to the consumption of stock holders), we introduce N productive assets. Each asset i produces output that evolves according to the stochastic process c_t^i , where $C_t = \sum_{i=1}^n c_t^i$. We model the output of each asset as the product of a systematic shock and an idiosyncratic shock:

$$c_t^i = Z_t \cdot \alpha_t^i, \quad \forall i, \quad (5)$$

where Z_t represents an economy-wide shock and α_t^i represents an individual asset’s sensitivity to the shock. This setup is similar to the setup of Bakshi and Chen (1997), who derive closed form solutions for bond and stock prices in an economy with standard power utility. Menzly, Santos and Veronesi (2004) formulate a continuous time economy extending the Campbell-Cochrane framework to multiple dividend processes.

Since the aggregate endowment equals the sum of individual asset outputs,

$$C_t = Z_t \cdot \sum_{i=1}^n \alpha_t^i. \quad (6)$$

Let D_t equal the output of one of the productive assets. Throughout the remainder of the paper, we shall interpret this asset as a particular equity market, such as the S&P 500. In such a case, D_t would equal the dividend process (broadly defined) for the S&P 500. Without loss of generality, let $D_t = c_t^1$.

Define $CD_t = \frac{1}{\alpha_t} \sum_{i=1}^n \alpha_t^i$. Thus, we may express aggregate consumption as:

$$\begin{aligned} C_t &= Z_t \cdot \sum_{i=1}^n \alpha_t^i \\ &= \alpha_t^1 Z_t CD_t \\ &= D_t CD_t \end{aligned} \tag{7}$$

Equation (7) expresses aggregate consumption as the product of the output of one asset and CD_t . CD_t can be written as the ratio of aggregate consumption to “dividends,” thus representing the remaining elements of aggregate consumption not represented by dividends. Such elements may include non-traded goods and real estate returns.

Finally, combining (2) and (7), we may write the pricing kernel as:

$$m_{t+1} = \ln(\beta) - \gamma \Delta d_{t+1} - \gamma \Delta cd_{t+1} + \gamma \Delta q_{t+1}, \tag{8}$$

where $m_{t+1} = \ln(M_{t+1})$, $\Delta d_{t+1} = \ln\left(\frac{D_{t+1}}{D_t}\right)$, and $\Delta cd_{t+1} = \ln\left(\frac{CD_{t+1}}{CD_t}\right)$.⁴

Because dividends and consumption are non-stationary we now must model dividend growth and the consumption-dividend ratio, CD_t . The main econometric issue is whether CD_t is stationary or, in other words, whether consumption and dividends are cointegrated. Table 1 reports some characteristics of the consumption-dividend ratio using total nondurables consumption and services as the consumption measure in addition to stationarity tests for CD_t . The first autocorrelation of the annual consumption dividend ratio is in the fairly high range of 0.86. When we test for a unit root in a specification allowing for a time trend and additional autocorrelation in the regression, we strongly reject the null hypothesis of a unit root. The test for the null hypothesis of no trend and a unit root only narrowly fails to reject at the 5% level. As a result we assume dividends and consumption are cointegrated with $[1, -1]$ as the cointegrating vector, and in our actual specification, we do allow for a time trend to capture the different means of consumption and dividend growth.

⁴The pricing kernel can also be obtained as the equilibrium implication of a production economy, as in Cox, Ingersoll, and Ross (1985) or in the analogous discrete-time version of Brock (1982). Assume that the representative agent possesses log utility. In addition, assume that there is one constant returns to scale technology, where Ω_{t+1} is the gross rate of return on the production opportunity from time t to time $t+1$. As demonstrated by Sun (1992) and Constantinides (1992), the pricing kernel is simply the inverse of Ω_{t+1} : $M_{t+1} = \frac{1}{\Omega_{t+1}}$. Therefore, for a particular specification of Ω_{t+1} , a specific pricing kernel is implied.

Our model for dividend growth and the consumption-dividend ratio becomes

$$\begin{aligned}\Delta d_{t+1} &= \mu_d + \rho_d \Delta d_t + \rho_{dq} q_t + \sigma_d \varepsilon_{t+1}^d + \kappa_1 \sqrt{q_t} \varepsilon_{t+1}^q \\ cd_{t+1} &= \lambda + \delta t + u_{t+1}\end{aligned}\tag{9}$$

$\mu_d, \rho_d, \rho_{dq}, \sigma_d$ and κ_1 are parameters governing dividend growth, Δd_t . The constant λ is without consequences once the model is put in stationary format, but the trend term, δt , accommodates different means for consumption growth and dividend growth. Specifically,

$$\Delta c_{t+1} = \delta + \Delta d_{t+1} + \Delta u_{t+1}\tag{10}$$

The model for u_{t+1} , the stochastic component of the consumption dividend ratio, is symmetric with the model for dividend growth:

$$u_{t+1} = \rho_u u_t + \rho_{uq} q_t + \sigma_{ud} \varepsilon_{t+1}^d + \sigma_u \varepsilon_{t+1}^u + \kappa_2 \sqrt{q_t} \varepsilon_{t+1}^q.\tag{11}$$

The conditional means of both dividend growth and the u_t process depend on their own lag and past preference shocks, q_t . This implies that dividend growth and the consumption dividend ratio are both ARMA(2,1) processes. Bansal and Yaron (2002) have recently stressed the importance of allowing an MA component in the dividend process.

We allow for two sources of correlation between dividend growth and the consumption dividend ratio, through σ_{ud} and through their joint dependence on preference shocks. The latter channel drives the conditional covariance between consumption growth and preference shocks. In particular, this covariance equals:

$$Cov_t [\Delta c_{t+1}, q_{t+1}] = (\kappa_1 + \kappa_2) \sigma_q q_t\tag{12}$$

Hence, if κ_1 and κ_2 are both negative, increases in q_t are associated with bad consumption growth shocks. This counter-cyclical risk aversion is imposed in the Campbell-Cochrane framework. In our framework, we leave these parameters unrestricted.

Another issue that arises in modeling consumption and stochastic risk aversion dynamics is whether the model preserves the notion of habit persistence. For this to be the case, even though consumption and risk aversion are negatively correlated, the habit stock should be a slowly decaying

moving average of past consumption. This is the case in this model but the relation is much more complex than in the univariate i.i.d. Campbell-Cochrane model, because of the presence of three autocorrelated stochastic variables driving the dynamics of consumption. Campbell and Cochrane (1999) also parameterize the process for the surplus ratio such that the derivative of the log of the habit stock is always positive with respect to log consumption. The habit stock in our model satisfies, $H_t = v_t C_t$, where $v_t = 1 - \frac{1}{Q_t}$ is in $(0,1)$ and is increasing in Q_t . That is, when risk aversion is high, the habit stock moves closer to the consumption level as is true in any habit model. It is now easy to see that the derivative condition above requires $\sigma_q / (\kappa_1 + \kappa_2) > 1 - Q_t$ for all t . Note that the right-hand side is negative and that even with negative κ_1 and κ_2 , this condition is not necessarily satisfied.

2.2.2 Inflation

One challenge with confronting consumption-based models with the data is that the model concepts have to be translated into nominal terms. Although inflation could play an important role in the relation between bond and stock returns, we want to assess how well we can match the salient features of the data without relying on intricate inflation dynamics and risk premiums. Therefore, we append the model with a simple inflation process:

$$\pi_{t+1} = \mu_\pi + \rho_\pi \pi_t + \sigma_\pi \varepsilon_{t+1}^\pi \quad (13)$$

Furthermore, we assume that the inflation shock is independent of all other shocks, in particular shocks to the real pricing kernel (or intertemporal marginal rate of substitution). The pricing of nominal assets then occurs with a nominal pricing kernel, \hat{m}_{t+1} that is a simple transformation of the real pricing kernel, m_{t+1} .

$$\hat{m}_{t+1} = m_{t+1} - \pi_{t+1} \quad (14)$$

2.3 The Full Model

We are now ready to present the full model. The logarithm of the pricing kernel or stochastic discount factor in this economy follows from the preference specification and is given by:

$$m_{t+1} = \ln(\beta) - \gamma \Delta c_{t+1} + \gamma \Delta q_{t+1} \quad (15)$$

Because of the logarithmic specification, the actual pricing kernel, M_{t+1} , is a positive stochastic process that ensures that all assets i are priced such that

$$1 = E_t [M_{t+1} (1 + R_{i,t+1})] \quad (16)$$

where $R_{i,t+1}$ is the percentage real return on asset i over the period from t to time $(t + 1)$, and E_t denotes the expectation conditional on the information at time t . Because M_t is strictly positive, our economy is arbitrage-free (see Harrison and Kreps (1979)). The model is completed by the specifications previously introduced of the fundamentals processes, which we collect here:

$$\begin{aligned} q_{t+1} &= \mu_q + \rho_q q_t + \sigma_q \sqrt{q_t} \varepsilon_{t+1}^q \\ \Delta d_{t+1} &= \mu_d + \rho_d \Delta d_t + \rho_{dq} q_t + \sigma_d \varepsilon_{t+1}^d + \kappa_1 \sqrt{q_t} \varepsilon_{t+1}^q \\ u_{t+1} &= \rho_u u_t + \rho_{uq} q_t + \sigma_{ud} \varepsilon_{t+1}^d + \sigma_u \varepsilon_{t+1}^u + \kappa_2 \sqrt{q_t} \varepsilon_{t+1}^q \\ \Delta c_{t+1} &= \delta + \Delta d_{t+1} + \Delta u_{t+1} \\ \pi_{t+1} &= \mu_\pi + \rho_\pi \pi_t + \sigma_\pi \varepsilon_{t+1}^\pi \end{aligned} \quad (17)$$

The real kernel process, m_{t+1} , is heteroskedastic, with its conditional variance moving with q_t . In particular,

$$Var_t [m_{t+1}] = \gamma^2 \left[(\sigma_d + \sigma_{ud})^2 + \sigma_u^2 + (\kappa_1 + \kappa_2 - \sigma_q)^2 \cdot q_t \right] \quad (18)$$

Consequently, increases in q_t will increase the Sharpe Ratio of all assets in the economy, and the effect will be greater the larger the quantity $|\kappa_1 + \kappa_2 - \sigma_q|$ is. If q_t and Δc_t are negatively correlated, the Sharpe Ratio of assets will increase during economic downturns (decreases in Δc_t).

3 Bond and Stock Pricing in the Moody Investor Economy

3.1 A General Pricing Model

We collect the state variables in the vector $Y_t = [q_t, \Delta d_t, u_t, \pi_t]'$. As shown in the Appendix, the dynamics of Y_t described in Equation (17) represent a simple, first-order vector autoregressive

process:

$$\begin{aligned} Y_t &= \mu + AY_{t-1} + (\Sigma_F F_{t-1} + \Sigma_H) \varepsilon_t \\ F_t &= (\parallel \phi + \Phi Y_t \parallel)' \odot I, \end{aligned} \tag{19}$$

where Y_t is the state vector of length k , μ and ϕ are parameter vectors also of length k and A , Σ_F , Σ_H and Φ are parameter matrices of size $(k \times k)$. $\varepsilon_t \sim N(0, I)$, I is the identity matrix of dimension k , $\parallel \cdot \parallel$ denotes the non-negativity operator for a vector⁵, and \odot denotes the Hadamard Product.⁶ Also, let the real pricing kernel be represented by:

$$m_{t+1} = \mu_m + \Gamma'_m Y_t + (\Sigma'_{mF} F_t + \Sigma'_{mH}) \varepsilon_{t+1}$$

where μ_m is a scalar and Γ_m , Σ_{mF} , and Σ_{mH} are k -vectors of parameters. We require the following restrictions:

$$\begin{aligned} \Sigma_F F_t \Sigma'_H &= 0 \\ \Sigma'_{mF} F_t \Sigma_{mH} &= 0 \\ \Sigma_H F_t \Sigma_{mF} &= 0 \\ \Sigma_F F_t \Sigma_{mH} &= 0 \\ \phi + \Phi Y_t &\geq 0 \end{aligned} \tag{20}$$

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 for some assets.

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 are only approximate in the event that the last restriction in Equation (20) is violated. If these variables are forced to reflect at zero, our use of the conditional lognormality features of the state variables becomes incorrect. It is for exactly this reason that we

⁵Specifically, if v is a k -vector, then $\parallel v \parallel = w$ where $w_i = \max(v_i, 0)$ for $i = 1, \dots, k$.

⁶The Hadamard Product operator denotes element-by-element multiplication. We define it formally in the Appendix. A useful implication of the Hadamard Product is that if $\phi + \Phi Y_t \geq 0$, for all elements, then $F_t F'_t = (\phi + \Phi Y_t) \odot I$.

model f directly and bound it from below in the specification of q_t in Equation (4), thus insuring that such instances are sufficiently rare.

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:

$$P_{n,t} = E_t [\exp(\hat{m}_{t+1}) P_{n-1,t+1}], \quad (21)$$

where $\hat{m}_{t+1} = m_{t+1} - \pi_{t+1}$ is the log of the nominal pricing kernel as argued above. 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 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:*

$$p_{n,t} = a_n^0 + a_n' Y_t \quad (22)$$

where the scalar a_n^0 and $(k \times 1)$ vector a_n' are defined recursively by the equations,

$$\begin{aligned} a_n^0 &= a_{n-1}^0 + (a_{n-1} - e_\pi)' \mu + \mu_m + \frac{1}{2} ((\Sigma_F' (a_{n-1} - e_\pi)) \odot (\Sigma_F' (a_{n-1} - e_\pi)))' \phi \\ &\quad + \frac{1}{2} (a_{n-1} - e_\pi)' \Sigma_H \Sigma_H' (a_{n-1} - e_\pi) + \frac{1}{2} (\Sigma_{mF} \odot \Sigma_{mF})' \phi \\ &\quad + \frac{1}{2} \Sigma_{mH}' \Sigma_{mH} + \frac{1}{2} \sigma_m^2 + (a_{n-1} - e_\pi)' [(\Sigma_{mF}' \odot \Sigma_F) \phi + \Sigma_H \Sigma_{mH}] \\ a_n' &= (a_{n-1} - e_\pi) A + \Gamma_m' + \frac{1}{2} ((\Sigma_F' (a_{n-1} - e_\pi)) \odot (\Sigma_F' (a_{n-1} - e_\pi)))' \Phi \\ &\quad + \frac{1}{2} (\Sigma_{mF} \odot \Sigma_{mF})' \Phi + (a_{n-1} - e_\pi) [(\Sigma_{mF}' \odot \Sigma_F) \Phi] \end{aligned} \quad (23)$$

and $a_0^0 = 0$ and $a_0' = -e_\pi$.

Notice that the log 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 models⁷. The process for the one-period short rate process, $r_t \equiv y_{1,t}$, is therefore simply $-(a_1^0 + a_1' Y_t)$.

Let $R_{n,t+1}^b$ and $r_{n,t+1}^b$ denote the nominal simple net return and log return, respectively, on an

⁷For an analysis of continuous time affine term structure models, see Dai and Singleton (2000).

n -period zero coupon bond between dates t and $t + 1$. Therefore:

$$\begin{aligned} R_{n,t+1}^b &= \exp(a_{n-1}^0 - a_n^0 + a'_{n-1}Y_{t+1} - a'_nY_t) - 1, \\ r_{n,t+1}^b &= a_{n-1}^0 - a_n^0 + a'_{n-1}Y_{t+1} - a'_nY_t. \end{aligned} \quad (24)$$

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:

$$V_t = E_t [\exp(m_{t+1}) (D_{t+1} + V_{t+1})]. \quad (25)$$

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

$$PD_t = \frac{V_t}{D_t} = E_t \left\{ \sum_{n=1}^{\infty} \exp \left[\sum_{j=1}^n (m_{t+j} + \Delta d_{t+j}) \right] \right\}, \quad (26)$$

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

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:*

$$PD_t = \sum_{n=1}^{\infty} \exp(b_n^0 + b'_n Y_t) \quad (27)$$

where the scalar b_n^0 and $(k \times 1)$ vector b'_n are defined recursively by the equations,

$$\begin{aligned} b_n^0 &= b_{n-1}^0 + (b_{n-1} + e_d)' \mu + \mu_m + \frac{1}{2} ((\Sigma'_F (b_{n-1} + e_d)) \odot (\Sigma'_F (b_{n-1} + e_d)))' \phi \\ &\quad + \frac{1}{2} (b_{n-1} + e_d)' \Sigma_H \Sigma'_H (b_{n-1} + e_d) + \frac{1}{2} (\Sigma_{mF} \odot \Sigma_{mF})' \phi \\ &\quad + \frac{1}{2} \Sigma'_{mH} \Sigma_{mH} + \frac{1}{2} \sigma_m^2 + (b_{n-1} + e_d)' [(\Sigma'_{mF} \odot \Sigma_F) \phi + \Sigma_H \Sigma_{mH}] \\ b'_n &= (b_{n-1} + e_d)' A + \Gamma'_m + \frac{1}{2} ((\Sigma'_F (b_{n-1} + e_d)) \odot (\Sigma'_F (b_{n-1} + e_d)))' \Phi \\ &\quad + \frac{1}{2} (\Sigma_{mF} \odot \Sigma_{mF})' \Phi + (b_{n-1} + e_d) [(\Sigma'_{mF} \odot \Sigma_F) \Phi] \end{aligned} \quad (28)$$

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:

$$\begin{aligned} R_{t+1}^e &= \exp(\pi_{t+1} + \Delta d_{t+1}) \left(\frac{\sum_{n=1}^{\infty} \exp(b_n^0 + b_n' Y_{t+1}) + 1}{\sum_{n=1}^{\infty} \exp(b_n^0 + b_n' Y_t)} \right) - 1 \\ r_{t+1}^e &= (\pi_{t+1} + \Delta d_{t+1}) + \ln \left(\frac{\sum_{n=1}^{\infty} \exp(b_n^0 + b_n' Y_{t+1}) + 1}{\sum_{n=1}^{\infty} \exp(b_n^0 + b_n' Y_t)} \right). \end{aligned} \quad (29)$$

The only intuition immediately apparent from comparing Equations (23) and (28) is that the coefficient recursions look identical except for the presence of the vector $-e_\pi$ in the bond equations and $+e_d$ in the equity equations. Because e_π selects inflation from the state variables, its presence accounts for the nominal value of the bond's cash flows with inflation depressing the bond price. Because e_d selects dividend growth from the state variables, its presence reflects the fact that equity is essentially a consol with real, stochastic coupons.

3.2 The risk free rate and the term structure

To obtain some intuition about the term structure, we start by calculating the log of the inverse of the conditional expectation of the (gross) pricing kernel, finding,

$$\begin{aligned} r_t^{\text{real}} &= -\ln(\beta) + \gamma(\mu_d + \delta - \mu_q) - \frac{1}{2}\gamma^2(\sigma_d + \sigma_{ud})^2 - \frac{1}{2}\gamma^2\sigma_u^2 + \\ &\quad \gamma\rho_d\Delta d_t + \gamma(\rho_u - 1)u_t + \left[\gamma(\rho_{dq} + \rho_{uq}) + \gamma(1 - \rho_q) - \frac{1}{2}\gamma^2(\kappa_1 + \kappa_2 - \sigma_q)^2 \right] q_t. \end{aligned} \quad (30)$$

Hence, the real interest rate follows a three-factor model with two observed factors (dividend growth and the consumption-dividend ratio) and one unobserved factor - a preference shock. It is useful to compare this to a standard version of the Lucas economy within which Mehra and Prescott (1985) documented the so-called low risk-free rate puzzle. The real risk free rate in the standard Mehra Prescott economy is given by

$$r_t^{\text{real,M-P}} = -\ln(\beta) + \gamma E_t(\Delta c_{t+1}) - \frac{1}{2}\gamma^2 V_t(\Delta c_{t+1}). \quad (31)$$

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. Because aggregate consumption growth exhibits quite low volatility, the latter term is typically of second-order importance.

The real rate in the Moody investor economy, r_t^{real} , equals the real rate in the Mehra-Prescott economy, plus three additional terms:

$$r_t^{\text{real}} = r_t^{\text{real,M-P}} + \gamma [(1 - \rho_q) q_t - \mu_q] - \frac{1}{2} \gamma^2 (\kappa_1 + \kappa_2 - \sigma_q)^2 q_t + \gamma (\rho_{dq} + \rho_{uq}) q_t \quad (32)$$

The first of the extra terms represents an additional consumption-smoothing effect. In this economy, risk aversion is also affected by q_t , and not only γ . When q_t is above its unconditional mean, $\mu_q/(1 - \rho_q)$, the consumption-smoothing effect is exacerbated. The second of the extra terms represents an additional precautionary savings effect. The uncertainty in stochastic risk aversion has to be hedged as well, depressing interest rates. The third term arises from the assumption that q_t affects the conditional mean of consumption growth. If it increases the conditional mean, the consumption smoothing effect is exacerbated. However, it is likely that ρ_{dq} and ρ_{uq} are negative since large preference shocks are associated with recessions. Taken together, these additional terms provide sufficient channels for this economy to avoid the risk-free rate puzzle.

In the data, we measure nominal interest rates. The nominal risk free interest rate in this economy simply follows from,

$$\exp(-r_t^f) = E_t [\exp(m_{t+1} - \pi_{t+1})]. \quad (33)$$

Because of the assumptions regarding the inflation process, the model yields an “approximate” version of the Fisher equation, where the approximation becomes more exact the lower the inflation volatility term.⁸:

$$r_t^f = r_t^{\text{real}} + \mu_\pi + \rho_\pi \pi_t - \frac{1}{2} \sigma_\pi^2. \quad (34)$$

⁸The expected gross ex-post real return on a nominal one-period contract, $E_t[\exp(r_t^f - \pi_{t+1})]$ will be exactly equal to the gross ex-ante real rate, $\exp(r_t^{\text{real}})$.

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.

Because of the neutrality of inflation, the model must generate an upward sloping term structure, a salient feature of term structure data, through the real term structure. To obtain some simple intuition about the determinants of the term spread, we investigate a two period real bond. For this bond, the term spread can be written as:

$$r_{t,2}^f - r_t^f = \frac{1}{2} \left(E_t \left[r_{t+1}^f \right] - r_t^f \right) + \frac{1}{2} Cov_t \left[m_{t+1}, r_{t+1}^f \right] - \frac{1}{4} Var_t \left[r_{t+1}^f \right] \quad (35)$$

The term in the middle determines the unconditional term premium, together with the third term, which is a Jensen's inequality term. The full model implies a quite complex expression for the unconditional term premium that cannot be signed. Under some simplifying assumptions, we can develop some intuition. First, we proceed under the assumption that the Jensen's inequality term is second order and can be ignored. Hence, we focus on the middle term. In general, we can write:

$$Cov_t \left[m_{t+1}, r_{t+1}^f \right] = v_0 + v_1 q_t \quad (36)$$

The time-variation in the term premium is entirely driven by stochastic risk aversion. Further assume that there is little persistence in the consumption dividend ratio and dividend growth and that there is no feedback from q_t to either (that is, ρ_{dq} and ρ_{uq} are both zero). In this case, $v_0 = 0$ and

$$v_1 = \gamma \theta (\sigma_q - \kappa_1 - \kappa_2) \quad (37)$$

with $\theta = \gamma (\rho_{dq} - \rho_{uq}) + \gamma (1 - \rho_q) - \frac{1}{2} \gamma^2 (\kappa_1 + \kappa_2 - \sigma_q)$. The interpretation is straightforward. The parameter θ measures whether the precautionary savings or consumption smoothing effect dominates in the determination of interest rates. Wachter (2002) also generalizes the Campbell – Cochrane setting to a two-factor model with one parameter governing the dominance of either one of these effects.

If $\theta > 0$, the consumption smoothing effect dominates and increases in q_t increase short rates. We see that this will also increase the term premium and give rise unconditionally to an upward sloping yield curve: bonds are risky in such a world. In contrast, when $\theta < 0$, the precautionary savings effect dominates. Increases in q_t now lower short rates, driving up the prices of bonds. Consequently,

bonds are good hedges against movements in q_t and do not require a positive risk premium.

3.3 Equity Pricing

In order to develop some intuition on the stock pricing equation in Equation (28), we split up the b_n vector into its four components. First, the component corresponding to inflation, denoted b_n^π , is zero because inflation is neutral in our model. Second, the coefficient multiplying current dividend growth is given by

$$b_n^d = (b_{n-1}^d + 1) \rho_d - \gamma \rho_d \quad (38)$$

This equation has the following interpretation. Dividends are the cash flows of the equity shares and therefore an increase in current dividends should raise future expected cash flows and hence the price-dividend ratio as long as dividend growth is positively autocorrelated. This is reflected by the term $(b_{n-1}^d + 1) \rho_d$. Additionally, dividend growth also helps determine consumption growth and therefore enters the pricing kernel (see the appendix for the exact formulation). Increases in dividends increase consumption and therefore lower marginal utility. The resulting increased discount rate depresses the price-dividend ratio. This effect is represented by the term, $-\gamma \rho_d$. In a standard Lucas-type model where consumption equals dividends and consumption growth is the only state variable, these are the only two effects affecting stock prices. Because they are countervailing effects, it is difficult to generate much variability in price dividend ratios in such a model.

In our model, consumption growth equals dividend growth plus the change in the consumption dividend ratio. This gives rise to an additional discount rate effect. Because the consumption dividend ratio is mean-reverting, an increase in u_t signals future expected decreases in consumption, and hence a lower current discount rate. Therefore, the effect of the consumption dividend ratio, b_n^u is given by,

$$b_n^u = b_{n-1}^u \rho_u + (1 - \rho_u) \gamma \quad (39)$$

Because ρ_u must be smaller than 1, increases in the consumption dividend ratio raise the price-dividend ratio. The first term, $b_{n-1}^u \rho_u$, accounts for the persistence in u_t . Finally, the price dividend ratio is affected by changes in risk aversion, q_t . The effect of q_t on the price dividend ratio is very

complex:

$$\begin{aligned}
b_n^q &= (b_{n-1}^d + 1) \rho_{dq} + b_{n-1}^u \rho_{uq} + b_{n-1}^q \rho_q \\
&\quad + \gamma (\rho_{dq} - \rho_{uq} + \rho_q - 1) + \frac{1}{2} \gamma^2 (\sigma_q - \kappa_1 - \kappa_2)^2 \\
&\quad + \frac{1}{2} (\sigma_q b_{n-1}^q + \kappa_1 (b_{n-1}^d) + \kappa_2 b_{n-1}^u)^2 \\
&\quad + \gamma (\sigma_q - \kappa_1 - \kappa_2) (\kappa_1 + \sigma_q b_{n-1}^q + \kappa_1 b_{n-1}^d + \kappa_2 b_{n-1}^u)
\end{aligned} \tag{40}$$

It is tempting to think that increases in risk aversion unambiguously depress price dividend ratios, but this is not necessarily true because q_t affects the price-dividend ratios through many channels. The first two terms on the right hand side of Equation (40) arise because q_t may change the conditional expectations of dividend growth and the consumption dividend ratio. If increases in q_t coincide with periods where dividend growth and the consumption dividend ratio are expected to be low (countercyclicity), these coefficients may be negative in which case shocks to q_t exacerbate negative cash flow effects. The third term accounts for the persistence in q_t ; the higher ρ_q is the more sensitive the price dividend ratio will be to changes in q_t . The fourth term summarizes the effect of q_t on the short rate. If the short rate increases when risk aversion increases, increases in risk aversion decrease price-dividend ratios, but if the precautionary savings motive dominates in the interest rate determination, we obtain the opposite effect. The fifth term arises from the fact that all conditional variances in this model are driven by q_t and this term is always positive.

The final term in Equation (40) is the most interesting as it arises from the covariance of the pricing kernel with q_t . It is larger in magnitude the larger is γ . The pricing effects totally depend on whether risk aversion is counter-cyclical or not. Let us define countercyclicity as having both κ_1 and κ_2 being less than zero. Let us also define the first-order effect through this channel as the terms not involving past b_{n-1} coefficients. If there is countercyclicity, consistent with intuition, increases in risk aversion through this sixth term have an unambiguous negative first-order effect on the price dividend ratio, of magnitude $\gamma (\sigma_q - \kappa_1 - \kappa_2) \kappa_1$. If the total initial effect of an increase in q_t on the price dividend ratio is indeed negative, the depressing effect of risk aversion increases on the price dividend ratio is exacerbated by terms involving b_{n-1}^q . Moreover, we already learned that, unless the consumption-dividend ratio is negatively autocorrelated, its initial effect on the price-dividend ratio is positive. This also implies a further depressing effect on the price dividend ratio when risk aversion increases coincide with consumption dividend ratio decreases, of magnitude $\kappa_2 b_{n-1}^u$. Hence,

the discount rate effect of a consumption dividend ratio decrease gets exacerbated by movements in stochastic risk aversion. Finally, if risk aversion increases in periods of poor cash flows, the cash flow effect is exacerbated by changes in risk aversion. Of course, a decrease in dividend growth has an offsetting negative effect on discount rates (see above), which makes the initial b_n^d coefficients possibly negative. If this happens, the discount rate effect of an increase in risk aversion is partially mitigated. Generally speaking, the effects of stochastic risk aversion on prices are rich and varied and if, for example, κ_1 and κ_2 were not to have the same sign, very complex dynamics are possible.

4 Estimation and Testing Procedure

4.1 Parameter Estimation

Our economy has four state variables, which we collected in the vector Y_t . Except for q_t , we assume that we can measure these variables from the data without error, with u_t being extracted from consumption-dividend ratio data. We are interested in the implications of the model for five endogenous variables: the short rate, r_t^f , the term spread, spd_t , the dividend yield, dp_t , the log excess equity return, r_t^{ex} , and the log excess bond return, r_t^{bx} . For all these variables we use rather standard data, comparable to what is used in the classic studies of Campbell and Shiller (1988) and Shiller and Beltratti (1992). Therefore, we describe the extraction of these variables out of the data and the data sources in a Data Appendix (Appendix A). We collect all the measurable variables of interest, the three observable state variables and the five endogenous variables in the vector W_t . Also, let Ψ denote the structural parameters of the model:

$$\Psi = [\mu_q, \mu_d, \mu_\pi, \delta, \rho_q, \rho_d, \rho_u, \rho_\pi, \rho_{dq}, \rho_{uq}, f, \sigma_d, \kappa_1, \sigma_{ud}, \sigma_u, \kappa_2, \sigma_\pi, \beta, \gamma]' \quad (41)$$

Throughout the estimation, we require that Ψ satisfies the conditions of Equations (20). There are a total of 19 parameters.

If we restrict ourselves to the term structure, the fact that the relation between endogenous variables and state variables is affine greatly simplifies the estimation of the parameters. As is apparent from Equations (27) and (29), the relationship between the dividend yield and excess equity returns and the state variables is non-linear. In the Computational Appendix, we linearize this relationship and show that the approximation is very accurate. Note that this approach is

very different from the popular Campbell-Shiller (1988) and Campbell (1990) linearization method, which linearizes the return expression itself before taking the linearized return equation through a present value model. We first find the correct solution for the price-dividend ratio and linearize the resulting equilibrium. The appendix demonstrates that the differences between the analytic and approximate moments do not affect our results.

Conditional on the linearization, the following property of W_t obtains,

$$W_t = \mu^w + \Gamma^w Y_{t-1} + (\Sigma_F^w F_{t-1} + \Sigma_H^w) \varepsilon_t \quad (42)$$

where the coefficients superscripted with ‘ w ’ are nonlinear functions of the model parameters, Ψ . Because Y_t follows a linear process with square-root volatility dynamics, unconditional moments of Y_t are available analytically as functions of the underlying parameter vector, Ψ . Let $X(W_t)$ be a vector valued function of W_t . For the current purpose, $X(\cdot)$ will be comprised of first and second order monomials, unconditional expectations of which are uncentered first and second moments of W_t . Using Equation (42), we can also derive the analytic solutions for uncentered moments of W_t as functions of Ψ . Specifically,

$$E[X(W_t)] = f(\Psi) \quad (43)$$

where $f(\cdot)$ is also a vector valued function (subsequent appendices provide the exact formulae). This immediately suggests a simple GMM based estimation strategy. The GMM moment conditions are,

$$g_{1T}(W_t; \Psi_0) = \frac{1}{T} \sum_{t=1}^T X(W_t) - f(\Psi_0). \quad (44)$$

Moreover, the additive separability of data and parameters in Equation (44) suggests a ‘fixed’ optimal GMM weighting matrix free from any particular parameter vector and based on the data alone. Specifically, the optimal GMM weighting matrix is the inverse of the spectral density at frequency zero of $g_{1T}(W_t; \Psi_0)$, which we denote as $S_{11}(W_T)$ because only the first term on the right hand side of Equation (44) contains any random variables (data). Concretely, we use the Newey West (1987) methodology with two ‘Newey-West’ lags. The inverse of $S_{11}(W_t)$ is the optimal weighting matrix. To estimate the system, we minimize the standard GMM objective function,

$$J(W_T; \hat{\Psi}) = g_{1T}(\hat{\Psi}) \cdot S_{11}^{-1}(W_t) \cdot g_{1T}(\hat{\Psi})' \quad (45)$$

in a one step optimal GMM procedure.

Because this system is extremely non-linear in the parameters, we took precautionary measures to assure that a global minimum has indeed been found. First, over 100 starting values for the parameter vector are chosen at random from within the parameter space. From each of these starting values, we conduct preliminary minimizations. We discard the runs for which estimation fails to converge, for instance, because the maximum number of iterations is exceeded, but retain converged parameter values as ‘candidate’ estimates. Next, each of these candidate parameter estimates is taken as a new starting point and minimization is repeated. This process is repeated for several rounds until a global minimizer has been identified as the parameter vector yielding the lowest value of the objective function. In this process, the use of a fixed weighting matrix is critical. Indeed, in the presence of a parameter-dependent weighting matrix, this search process would not be well defined. Finally, the parameter estimates producing the global minimum are confirmed by starting the minimization routine at small perturbations around the parameter estimate, and verifying that the routine returns to the global minimum.

4.2 Moment Conditions

We use a total of 28 moment conditions to estimate the model parameters. They can be ordered into 4 groups. The first set is simply the unconditional means of the W_t variables:

$$E \left[\Delta d_t, \Delta c_t, \pi_t, r_t^f, dp_t, spd_t, r_t^{ex}, r_t^{bx} \right]$$

In this group only the mean of the spread and the excess bond return are moments that could not be investigated in the original Campbell – Cochrane framework. We also require the model to match the equity premium.

The second group includes the second moments of the state variables and the short rate and the correlation between short rates and the fundamentals processes,

$$E \left[(\Delta d_t)^2, (\Delta c_t)^2, (\pi_t)^2, \left(r_t^f \right)^2, (\Delta d_t \cdot \Delta c_t), \left(\Delta d_t \cdot r_t^f \right), \left(\Delta c_t \cdot r_t^f \right), \left(\pi_t \cdot r_t^f \right) \right] \quad (46)$$

This is a dimension that is absent in Campbell and Cochrane. By assuming a constant interest rate, they implicitly assume a zero correlation between short rates and fundamentals. We will see

later that this in itself is not such an important omission as the empirical movements are indeed close to zero. A third set of moments concerns the second moments of the endogenous variables:

$$E \left[(dp_t)^2, (spd_t)^2, (r_t^{ex})^2, (r_t^{bx})^2, (\Delta d_t \cdot r_t^{ex}), (\Delta c_t \cdot r_t^{ex}), (\Delta d_t \cdot r_t^{bx}), (\Delta c_t \cdot r_t^{bx}) \right] \quad (47)$$

This includes the volatility of both the dividend yield and excess equity returns, so that the estimation incorporates the excess volatility puzzle and adds to that the volatility of the term spread and bond returns. Intuitively, this may be a hard trade-off (see, for instance Bekaert (1996)). To match the volatility of equity returns and price dividend ratios, volatile intertemporal marginal rates of substitution are necessary, but interest rates are relatively smooth and bond returns are much less variable than equity returns in the data. They are functions of expected marginal rates of substitution and their variability must not be excessively high to yield realistic predictions. Finally; note that we also include the fundamentals on the one hand (consumption and dividend growth) and asset returns on the other hand (bond and stock returns). Therefore we confront the model directly with the Cochrane-Hansen (1990) puzzle.

The final set of moments simply ensures that the model approximately matches the persistence of the fundamental processes and the short rate:

$$E \left[(\Delta d_t \cdot \Delta d_{t-1}), (\Delta c_t \cdot \Delta c_{t-1}), (\pi_t \cdot \pi_{t-1}), (r_t^f \cdot r_{t-1}^f) \right] \quad (48)$$

Note that there are nine over-identifying restrictions and that we can use the standard J -test to test the fit of the model.

4.3 Tests of Extra Moments

If the model can fit the base moments, it would be a rather successful stock and bond pricing model. Nevertheless, we want to use our framework to fully explore the implications of a model with stochastic risk aversion for the joint dynamics of bond and stock returns, partially also to guide future research. In section 6, we consider a set of additional moment restrictions that we would like to test.

To test conformity of the estimated model with moments not explicitly fit in the estimation stage, the following GMM based statistic is constructed. Let $g_{2T}(\Psi_0, W_t)$ be the sample mean of

the restrictions we wish to test. By the Mean Value Theorem,

$$g_{2T}(\hat{\Psi}) \stackrel{a.s.}{=} g_{2T}(\Psi_0) + D_{2T} \cdot (\hat{\Psi} - \Psi_0) \quad (49)$$

where $D_{2T} = \frac{\partial g_{2T}(\Psi)}{\partial \Psi}$. Since $\hat{\Psi}$ is estimated from the first set of orthogonality conditions,

$$(\hat{\Psi} - \Psi_0) \stackrel{a.s.}{=} -(A_{11}D_{1T})^{-1} A_{11}g_{1T}(\Psi_0) \quad (50)$$

with

$$D_{1T} = \frac{\partial g_{1T}(\Psi_0)}{\partial \Psi} \quad (51)$$

$$A_{11} = D'_{1T}S_{11}^{-1} \quad (52)$$

Substituting,

$$g_{2T}(\hat{\Psi}) \stackrel{a.s.}{=} Lg_T(\Psi_0) \quad (53)$$

where

$$L = \left[-D_{2T} \cdot (A_{11}D_{1T})^{-1} A'_{11}, I \right] \quad (54)$$

$$g_T(\Psi_0) = \left[g_{1T}(\Psi_0)', g_{2T}(W_t; \Psi_0)' \right]' \quad (55)$$

Since $\sqrt{T}g_T(\Psi_0) \rightarrow N(0, S)$ where S is the spectral density at frequency zero of all the orthogonality conditions, and S_{11} is the top left quadrant of S , the statistic,

$$Tg_{2T}(\hat{\Psi})[LSL']^{-1}g_{2T}(\hat{\Psi}) \quad (56)$$

has a $\chi^2(k)$ distribution under the null, where k is the number of moments considered in $g_{2T}(\hat{\Psi})$.

5 Estimation Results

5.1 Parameters

Table 2 reports the parameter estimates for the model, using aggregate consumption of services and non-durables as the consumption measure. The first column reports mean parameters. The negative estimate for δ ensures that average consumption growth is lower than average dividend growth, as is true in the data. Importantly, μ_q is not estimated, but fixed at unity, reducing the number of estimated parameters to 18. Because risk aversion under this model is proportional to $\exp(q_t)$, the unconditional mean and volatility of q_t are difficult to jointly identify under the lognormal specification of the model. Restricting μ_q to be unity does not significantly reduce the flexibility of the model. The second column reports feedback coefficients. Dividend growth shows little serial persistence and the coefficient is negative. The consumption-dividend ratio is very persistent. Stochastic risk aversion has little effect on the conditional means of either. If we set these coefficients (ρ_{dq} and ρ_{uq}) equal to zero, the implied process for consumption growth is still an ARMA(2,2). The autoregressive coefficients would be $\rho_d + \rho_u = 0.895$ and $-\rho_d \cdot \rho_u = 0.0995$ respectively.

The volatility parameters are reported in the third column. The homoskedastic component of the correlation between dividend growth and the consumption dividend ratio is negative (σ_{ud}). As expected, the innovation to the consumption dividend ratio loads negatively on shocks to q_t , but the dividend innovation loads positively. Overall, shocks to q_t lower consumption growth because $\kappa_1 + \kappa_2 < 0$. Finally, we report the discount factor β and the curvature parameter of the utility function, γ . Because risk aversion is equal to γQ_t , and the economy is growing, these coefficients are difficult to interpret by themselves.

We also report the test of the over-identifying restrictions. There are 18 parameters and 28 moment conditions, making the J-test a $\chi^2(10)$ under the null. The test fails to reject at the 5% level, but would reject at the 10% level.

5.2 Implied Moments

Whereas the test suggests that the model provides an overall satisfactory fit with the data, we want to assess which moments it fits well and which moments it fails to fit perfectly. Table 3 shows a large array of first and second moments regarding fundamentals (dividend growth, consumption

growth and inflation), and endogenous variables (the risk free rate, the dividend yield, the term spread, excess equity returns and excess bond returns). We show the means, volatilities, first-order autocorrelation and the full correlation matrix. Numbers in parentheses are GMM based standard errors for the sample moments. Numbers in brackets are population moments for the model (using the log-linear approximation for the price dividend ratio described above for dp_t and r_t^{ex}). In our discussion, we informally compare sample with population moments using the data standard errors as a guide to assess goodness of fit. This of course ignores the sampling uncertainty in the parameter estimates.

5.2.1 The equity premium and risk free rate puzzles

Table 3 indicates that our model implies an excess return premium of 6% on equity, which matches the data moment of 5.9% almost exactly. We do not do so at the cost of exorbitantly high-risk free rates, as the average nominal interest rate is only 5.1%, which is just within two standard errors of the sample mean. Hence, our economy does not suffer from the risk free rate puzzle (see Weil 1989)). At the same time, we also generate an average excess bond return of 1.8%, which is higher than the 1% observed in the data, but again our model moment is comfortably within a two standard error bound.

5.2.2 Excess volatility

Stock returns are not excessively volatile from the perspective of our model. While the standard deviation of excess returns in the data is 19.7%, we generate excess return volatility of 17.6%. To accomplish this, the intertemporal marginal rate of substitution must be rather volatile in our model, and that often has the implication of making bond returns excessively volatile (see, for example, Bekaert (1996)). This also does not happen in the model, which generates an excess bond return volatility of 8.5% versus 8.1% in the data.

Table 4 helps interpret these results. It provides variance decompositions for a number of endogenous variables in terms of current and lagged realizations of the four state variables. The state variables are elements of Y_t , defined above.

Less than 10% of the variation in excess stock returns is explained by dividend growth and the consumption dividend ratio. Whereas stocks are a claim to real dividends and inflation is neutral in our model, we investigate nominal stock returns with nominal dividends and this explains the

0.55% of stock return variation attributable to inflation variation. The bulk of the variance of returns (over 90%) is explained by stochastic risk aversion. In Campbell and Cochrane (1999), this proportion is 100% because consumption and dividend growth are modeled as i.i.d. processes. Whereas the Campbell and Cochrane (1999) model featured a non-stochastic term structure, we are able to generate much variability in bond returns simply using a stochastic inflation process, which accounts for 75% of the variation in bond returns. The remainder is primarily due to stochastic risk aversion, and almost no variation is due to dividend growth or the consumption dividend ratio. This is consistent with the lack of a strong relationship between bond returns and these variables in the data.⁹

The excess volatility puzzle often refers to the inability of present value models to generate variable price-dividend ratios or dividend yields (see Campbell and Shiller (1988), Cochrane (1992)). In models with constant excess discount rates, price-dividend ratios must either predict future dividend growth or future interest rates and it is unlikely that predictable dividend growth or interest rates can fully account for the variation of dividend yields (see Ang and Bekaert (2003a) and Lettau and Ludvigson (2003) for recent articles on this topic). Table 3 shows that our model undershoots the variance of dividend yields (1.2% versus 1.4%) by very little. Table 4 shows that the bulk of this variation comes from stochastic risk aversion and not from cash flows. The 3% explained by the consumption dividend growth ratio is in effect also a discount rate effect, as we indeed model stocks as a claim to dividends and not consumption.

5.2.3 Term structure dynamics

One of the main goals of this article is to develop an economy that matches salient features of equity returns as in Campbell and Cochrane, while introducing a stochastic but tractable term structure model. Table 3 reports how well the model performs with respect to the short rate and the term spread. The volatilities of both are matched near perfectly. The model also reproduces a persistent short rate process, with an autocorrelation coefficient of 0.83 (versus 0.90 in the data). However, the term spread is not persistent enough with an autocorrelation coefficient of 0.486 versus 0.734 in the data, which constitutes a difference of 2.75 standard errors.

The term structure model in this economy is affine and variation in yields is driven by four

⁹In this data sample, a regression of excess bond returns on lagged real dividend growth and the lagged consumption dividend ratio yields an adjusted r-squared statistic of 0.03. A regression run on simulated data under the model reported in Table 2 yields an adjusted r-squared of under 0.01.

factors: dividend growth, the consumption-dividend ratio, inflation and stochastic risk aversion. Table 4 shows how much of the variation of the short rate and the term spread each of these factors explains. Interestingly, inflation shocks drive about 86.5% of the total variation of the short rate, but only 50.5% of the variation in the term spread. This is of course not surprising since the first-order effect of expected inflation shocks is to increase interest rates along the entire yield curve. Because there is no inflation risk premium in this model, the spread actually reacts negatively to a positive inflation shock, as inflation is a mean reverting process.

Whereas stochastic risk aversion only explains 3% of the variation in short rates, it drives 9.3% of the variation in the term spread. This is easy to understand as the term premium is entirely driven by variation in risk aversion. The term premium is proportional to the expected excess bond return in any term structure model and Table 4 confirms the fact that risk premiums in this model are naturally driven by stochastic risk aversion. We would expect that consumption growth would also explain some portion of the variability of the term structure (see Equation 30). In this model consumption growth is split up in the consumption-dividend ratio and dividend growth and it is the latter that mostly drives the variation in the term structure. Whereas dividend growth explains about 9% of the variation in short rates, it explains about 40% of the variation in term spreads which is, not surprisingly, counterfactual.¹⁰

Table 3 shows that in our economy negative consumption or dividend growth shocks (recessions) are associated with higher nominal short rates and lower spreads. Countercyclical interest rates and pro-cyclical term spreads are exactly what we observe in the US economy (see, for instance, Ang and Bekaert (2003b) for more detail), but the effects as measured relative to annual dividend and consumption growth are not very strong in the current data sample. For example, the unconditional correlation between dividend growth and interest rates is only slightly negative (-0.01) and indistinguishable from zero in the sample data, and the correlation between aggregate consumption growth and interest rates is slightly positive (0.01). In the model, both of these correlations are negative. The model generated correlation between the short rate and dividend growth is -0.28, just slightly more than two standard deviations below the sample statistic, and the model generated correlation between the short rate and consumption growth is -0.16, just more than one standard deviation below the sample estimate. The correlation between the term spread and both consumption

¹⁰Regressions of the short rate and yield spread on contemporaneous dividend and consumption growth both yield r-squared statistics of less than 0.01 in the data. Using simulated data based on the model in Table 2, these r-squared statistics are 0.09 and 0.49 respectively.

and dividend growth is negative in the sample, contrary to the pro-cyclicality story. Consequently, as Table 3 indicates, our model does not match these correlations particularly well, as the model correlations are positive and more than two standard deviations above their sample counterparts.

To help us interpret these findings, we determine the coefficients in the affine relation of the (nominal) short rate with the factors. We find that the short rate reacts positively to increases in risk aversion indicating that with respect to the q -shock the consumption smoothing effect dominates (see Equation (30)). However, the short rate reacts negatively to consumption growth shocks primarily because dividend growth is slightly negatively correlated. The term spread reacts more strongly to preference shocks and also reacts positively to an increase in dividend growth. Whereas these findings do not constitute strong evidence against the model, they do suggest an additional factor separating bond from stock return behavior would be beneficial.

5.2.4 Link between fundamentals and asset returns

Equilibrium models typically imply that consumption growth and stock returns are highly correlated. In our model, several channels break the tight link between stock returns and consumption growth in standard models. First, we model equity correctly as a claim to dividends, not to consumption. This helps reduce the correlation between stock returns and consumption growth, but it generates another puzzle. For equity to earn a risk premium, its returns must be correlated with the pricing kernel, and dividend growth and consumption growth are reportedly not highly correlated (see Campbell and Cochrane (1999)). Our second mechanism to break the tight link between consumption growth and stock returns comes in play here as well: stochastic risk aversion is the main driver of the variability of the pricing kernel.

Table 3 shows that the fit of our model with respect to the links between fundamentals and asset returns is phenomenal. First, aggregate consumption growth and dividend growth have a realistic 0.5 correlation in our model. Second, we generate a correlation between both dividend and consumption growth and excess equity returns of 0.21, which is not significantly above the correlation in the data (which is respectively 0.11 for dividend growth and 0.18 for consumption growth). Third, in the data, bond returns and dividend and consumption growth are negatively correlated but the correlation is small. We generate small correlations in the model, matching the sign in the case of dividend growth.

One interesting feature of our parameter estimates is that they fail to match the volatility of the

fundamentals. That is, although they are close to the two standard error bound around the sample estimate, the volatilities of dividend growth and consumption growth are both lower than they are in the data. This makes the great fit of our model with the data moments even more of a pleasant surprise. The main reason is that stochastic risk aversion, as in Campbell and Cochrane (1999), is instrumental in generating many salient features of the data.

5.2.5 Time-varying Risk Appetites

Stochastic risk aversion in our model equals γQ_t . Because it is unobserved, we characterize its properties through simulation in Table 5. Median risk aversion equals 2.69 and its interquartile range is [1.79, 5.05]. Risk aversion is positively skewed and the 90 percentile observation equals 11.3. It has less than a 1% chance of reaching 100. Stochastic risk aversion is quite persistent with a persistence parameter of 0.73

The bottom panel of Table 5 reports the correlation of risk aversion with all of our endogenous and exogenous variables. As expected, the variable is countercyclical, showing a negative correlation with both dividend and consumption growth. When risk aversion is high, dividend yields increase (that is, price dividend ratios decrease) making the dividend yield-risk aversion correlation positive. From Table 4, we already know that q_t is the sole driver of time-variation in risk and expected return in this model, driving up expected returns on both stocks and bonds in times of high-risk aversion. Therefore, periods of high risk aversion are characterized by negative realizations of unexpected returns as well as increased positive expected returns. The net unconditional correlation between risk aversion and returns is indeterminate and tends to be near zero.

6 The Joint Dynamics of Bond and Stock Returns

The economy we have created so far manages to match more salient features of the data than the original Campbell-Cochrane (1999) article in an essentially linear framework. Nevertheless, the main goal of this article is to ascertain how many of the salient features of the joint dynamics of bond and stock returns it can capture. In this section, we first analyze the comovements of bond and stock returns and then look at the predictability of bond and stock returns.

6.1 The bond-stock returns correlation

Shiller and Beltratti (1992) show that in a present value framework with constant risk premiums, the correlation between bond and stock returns is too low relative to the correlation in the data. Nevertheless, Table 3 shows that the correlation between bond and stock returns during the sample is only 0.152 with a relatively large standard error. In our model, expected excess bond and stock returns both depend negatively on stochastic risk aversion and this common source of variation induces additional correlation between bond and stock returns. We generate a correlation of 0.391, which is slightly more than two standard errors above the sample moment. A formal test, incorporating both sampling error and parameter uncertainty (see Section 4.3) is reported in Table 6, Panel A. The test rejects at the 5% but not at the 1% level. In this test, we do not incorporate the uncertainty of estimating the means and variances because these have been tested before. If we would do so, the test would not even reject at the 10% level.

The reason we overshoot the correlation has to do with the effect of q_t on asset prices. While equity prices decrease when q_t increases, the effect on bond returns is ambiguous because the effect of q_t on interest rates is ambiguous. Empirically, we have shown that q_t increases interest rates and hence lowers bond returns. Therefore, q_t provides a channel for higher correlation, both between expected and unexpected bond and stock returns.

Clearly, this is one dimension that could be a useful yardstick for future models. Dai (2003) argues that the bond market requires a separate factor that does not affect stock prices. We believe that it might be more fruitful to think about potential stochastic components in cash flows that are not relevant for bond pricing and that our simple dividend growth model may have missed. Another fruitful avenue for extending the model is to investigate the dynamics of inflation more closely. Campbell and Ammer (1993) empirically decompose bond and stock return movements into various components and find inflation shocks to be an important source of negative correlation between bond and stock returns.

6.2 Bond and Stock Return Predictability

Table 6, Panel B reports predictability tests. We run univariate regressions of excess bond and stock returns using four instruments: the risk free rate, the dividend yield, the yield spread, and the excess dividend yield (the dividend yield minus the interest rate). A long list of articles has

demonstrated the predictive power of these instruments for excess equity returns. However, a more recent literature casts doubt on the predictive power of the dividend yield, while confirming strong predictability for equity returns using the interest rate or term spread as a predictor, at least in post-1954 data, see for example Ang and Bekaert (2003a) and Campbell and Yogo (2002).

In our annual data set, the only significant predictor of equity returns is the yield spread. The t-statistics in the short rate and excess dividend yield regression are above 1.00 but do not yield a 5% rejection. When we investigate bond returns, we also find the yield spread to be the only significant predictor. This predictability reflects the well-known deviations from the Expectations Hypothesis (see Campbell and Shiller (1991) and Bekaert, Hodrick and Marshall (2001)). A higher yield spread predicts high expected excess returns on both stocks and bonds.

The predictability coefficients implied by our model are reported in parentheses. Of course, except for the yield spread regression, these tests have little power and are not useful to investigate. What is interesting is to check whether the model gets the signs right. The one miss here is the negative sign of the short rate coefficient in the equity return regression. This puzzle, more prevalent with post Treasury accord data and known since Fama and Schwert (1979), can potentially be resolved in our model, because the equity premium increases when risk aversion (q_t) increases, whereas the short rate can increase or decrease with higher risk aversion depending on whether the consumption smoothing or precautionary savings effect dominates. Wachter (2002) investigates a two-factor extension of Campbell and Cochrane (1999) with exactly this purpose. Because at our estimated values, an increase in q_t increases the short rate, we generate a positive correlation between current interest rates and the equity premium. A full investigation of this puzzle requires a more serious investigation of inflation dynamics, because the empirical relationship may be due to the expected inflation component in nominal interest rates, rather than the real short rate component.

With respect to the predictive power of the yield spread, the model does reasonably well. It generates substantial positive predictability coefficients in both cases and also matches the fact that the coefficient is larger for the equity than for the bond return regression. The formal tests fail to reject at the 5% level but would reject at the 10% level, because the coefficients are somewhat too small relative to the sample estimates. The reason the model generates this empirical fact has in fact already been discussed. In our section on the term structure, we reported that the spread reacts positively to changes in risk aversion, and risk premiums generally increase when q_t increases. Hence, the effect is immediate. In a recent paper, Buraschi and Jiltsov (2004) introduce external

habit in a monetary term structure model and examine the implied term structure dynamics. They also find that the model fits the deviations of the Expectations Hypothesis.

One strong implication of the model is that the predictable components in the excess returns of stocks and bonds are perfectly correlated because of the dependence on q_t . In the data, this would be the case if the yield spread was really the only true predictor. To investigate how realistic this implication of the model is, we project the excess returns in the data onto the interest rate, the yield spread and the dividend yield and compute the correlation of the two fitted values. We find this correlation to be 0.81 with a standard error of 0.29. This suggests the assumption of perfect correlation between expected excess returns on bonds and stocks is a rather accurate approximation of the truth.

7 Conclusion

In this article, we have presented a pricing model for stocks and bonds where potentially counter-cyclical preference shocks generate time-variation in risk premiums. The model can be interpreted as a tractable version of the external habit model of Campbell and Cochrane (1999) accommodating a fully stochastic term structure. Our fundamentals include both consumption (which enters the utility function) and dividends (which is the relevant cash flow process), which are assumed to be cointegrated processes.

A GMM estimation reveals that the model cannot be rejected with respect to its fit of a large number of salient features of the data, including the level and variability of interest rates, bond and stock returns, term spreads and dividend yields. The model also matches the correlation between fundamentals (consumption and dividend growth) and asset returns. We further examine the fit of the model with respect to bond and stock return dynamics, finding that it produces a somewhat too high correlation between stock and bond returns but matches the fact that the term spread signals high risk premiums on both. The model also does not generate a negative relation between the equity premium and short rates, although it could theoretically match this relation. This relationship deserves further scrutiny in a model where the inflation process gets more attention.

After earlier versions of this were circulated, several authors have also considered the joint modeling of bond and stock returns in similar but distinct frameworks. Wachter (2002) provides a two factor extension of Campbell and Cochrane (1999) model by making expected consumption growth

stochastic. Brandt and Wang (2003) model risk aversion as a function of unexpected inflation. Dai (2003) formulates a model with habit formation that nests both internal habit formation and external habit formation models. These articles empirically focus primarily on deviations from the expectations hypothesis. Mamaysky (2002) investigates the joint pricing of bond and stock returns in a continuous time model where the dividend yield is exogenous. Li (2002) combines a simplified version of our model with Mamaysky's to study the correlation of bond and stock returns. Our research reveals that future modeling efforts must search for factors that drive a stronger wedge between bond and stock pricing. Possible candidates are more intricate modeling of the inflation process and a cash flow component uncorrelated with the discount rate.

References

- [1] Abel, A., 1990, Asset prices under habit formation and catching up with the Joneses, *American Economic Review* 80, 38-42.
- [2] Abel, A., 1999, Risk premia and term premia in general equilibrium, *Journal of Monetary Economics* 43, 3-33.
- [3] Ait-Sahalia, Y., Parker, A. and M. Yogo, 2003, Luxury Goods and the Equity Premium, working paper, Princeton University.
- [4] Ang, A. and G. Bekaert, 2003a, Stock Return Predictability: Is it There?, working paper, Columbia University.
- [5] Ang, A. and G. Bekaert, 2003b, The Term Structure of Real Rates and Expected Inflation, working paper, Columbia University.
- [6] Ameriks, J. and S. Zeldes, 2001, Do Household Portfolio Shares Vary with Age?, working paper, Columbia University.
- [7] Bakshi, G. and Z. Chen, 1997, An Alternative Valuation Model for Contingent Claims, *Journal of Financial Economics*, 44, 123-165.
- [8] Bansal, R., and C. Lundblad, 2002, Marker Efficiency, Asset Returns, and the Size of the Risk Premium in Global Equity Markets, *Journal of Econometrics*, 109, 195-237.
- [9] Bansal, R. C. Lundblad and R. Dittmar, 2004, Consumption, Dividends and the Cross Section of Equity Returns, working paper, University of Michigan.
- [10] Bansal, R. and M. Yaron, 2001, Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles, *Journal of Finance*, forthcoming.
- [11] 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.
- [12] 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.

- [13] Bekaert, G., R. J. Hodrick and D. A. Marshall, 2001, Peso problem explanations for term structure anomalies, *Journal of Monetary Economics* 48, 241-270.
- [14] Brandt, M. and K. Q. Wang, 2003, Time-Varying Risk Aversion and Unexpected Inflation, *Journal of Monetary Economics*, 50, 1457-1498.
- [15] Brock, W., 1982, *Asset Prices in a Production Economy*, Scientific Essays of William Allen Brock, Elgar.
- [16] Buraschi, A. and A. Jiltsov, 2004, Habit Persistence and the Nominal Term Structure of Interest Rates, working paper, London Business School.
- [17] Campbell, J. Y., 1990, Measuring the Persistence of Expected Returns, *American Economic Review* 80, 43-47.
- [18] Campbell, J. Y., 1993, Intertemporal Asset Pricing without Consumption Data, *American Economic Review* 83, 487-512.
- [19] Campbell, J. Y., and J. Ammer, 1993, What moves the Stock and Bond Markets? A Variance Decomposition for Long-Term Asset Returns, *Journal of Finance*, 48, 3-37.
- [20] Campbell, J. Y., and Cochrane, J. H., 1995, By force of habit: A consumption based explanation of aggregate stock market behavior, working paper, NBER.
- [21] 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.
- [22] 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.
- [23] 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.
- [24] Campbell, J. Y. and M. Yogo, 2003, Efficient Tests of Stock Return Predictability, working paper, Harvard University.
- [25] Cochrane, J. H., 1992, Explaining the variance of price-dividend ratios, *The Review of Financial Studies* 5, 243-280.

- [26] Cochrane, J. H., and Hansen, L. P., 1992, Asset Pricing Explorations for Macroeconomics, NBER Macroeconomics Annual, 1992, 115-165
- [27] Constantinides, G. M., 1992, A theory of the nominal term structure of interest rates, Review of Financial Studies 5, 531-552.
- [28] Cox, J. C., J. E. Ingersoll and S. A. Ross, 1985, A theory of the term structure of interest rates, Econometrica 53, 385-408.
- [29] Dai, Q., 2003, Term Structure Dynamics in a Model with Stochastic Internal Habit, working paper.
- [30] Dai, Q. and K. Singleton, 2000, Specification Analysis of Affine Term Structure Models, Journal of Finance, 55, 1943-1978.
- [31] Dickey, D., and A. Fuller, 1984, Testing for Unit Roots: 2, Econometrica, 52, 1241-1269.
- [32] Fama, E., and K. French, 1989, Business conditions and expected returns on stocks and bonds, Journal of Financial Economics 25, 23-49.
- [33] Fama, E., and Schwert, 1977, Asset Returns and Inflation, Journal of Financial Economics, 5, 115-146.
- [34] Hamilton, J. D., 1994, Time Series Analysis, Princeton University Press.
- [35] Hansen, L. P., 1982, Large sample properties of generalized method of moments estimators, Econometrica 50, 1029-1054.
- [36] Hansen, L. P., and R. Jagannathan, 1991, Implications of security market data for models of dynamic economies, Journal of Political Economy 99, 225-262.
- [37] Harrison, J. M., and D. Kreps, 1979, Martingales and arbitrage in multi-period securities markets, Journal of Economic Theory 20, 381-408.
- [38] Heaton, J., 1995, An empirical investigation of asset pricing with temporally dependent preference specifications, Econometrica 63, 681-717.
- [39] Keim, D. B., and R. F. Stambaugh, 1986, Predicting returns in bond and stock markets, Journal of Financial Economics 17, 357-390.

- [40] Kumar, S., and A. Persaud, 2002, Pure Contagion and Investors' Shifting Risk Appetite, *International Finance*, 5, 401-436.
- [41] Lettau, M. and S. Ludvigson, 2004, Expected Returns and Expected Dividend Growth, working paper, New York University.
- [42] Lettau, M. and J. Wachter, Why is Long-Horizon Equity Less Risky? A Duration-Based Explanation of the Value Premium, working paper, New York University.
- [43] Lucas, R. E., Jr., 1978, Asset prices in an exchange economy, *Econometrica* 46, 1426-1446.
- [44] Lustig, H. and S. Van Nieuwerburgh, 2003, Housing Collateral, Consumption Insurance and Risk Premia: an Empirical Perspective, working paper, University of Chicago.
- [45] Mamaysky, H., 2002, Market Prices of Risk and Return Predictability in a Joint Stock-Bond Pricing model, Working Paper, Yale.
- [46] Mankiw, G. and S. Zeldes, 1991, The Consumption of Stockholders and Nonstockholders, *Journal of Financial Economics*, 29, 97-112.
- [47] Mehra, R., and E. C. Prescott, 1985, The equity premium: A puzzle, *Journal of Monetary Economics* 15, 145-161.
- [48] Menzly, L., Santos, T., and P. Veronesi, 2004, Understanding Predictability, *Journal of Political Economy*, 112, 1-47.
- [49] Newey, W., and K. West, 1987, A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica*, 55, 703-708.
- [50] Piazzesi, M., Schneider, M., and S. Tuzel, 2003, Housing, Consumption and Asset Pricing, working paper, University of Chicago.
- [51] Santos, T., and P. Veronesi, 2004, Labor Income and Predictable Stock Returns, working paper, University of Chicago.
- [52] 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.

- [53] Shiller, R.J., and A. Beltratti, 1992, Stock Prices and Bond Yields: Can Their Movements Be Explained in Terms of Present Value models?, *Journal of Monetary Economics*, 30, 25-46.
- [54] Sun, T., 1992, Real and Nominal Interest Rates: A Discrete-Time Model and Its Continuous Time Limit, *Review of Financial Studies*, 5, 581-611.
- [55] Wachter, J., 2002, Habit Formation and Returns on Stocks and Bonds, Working Paper, New York University.
- [56] Wei, M., 2003, Human Capital, Business Cycles, and Asset Pricing, working paper, Columbia University.
- [57] Weil, P., 1989, The equity premium puzzle and the risk free rate puzzle, *Journal of Monetary Economics* 24, 401-421.

Table 1: Consumption - Dividend Ratio Characteristics

Univariate Statistics			ADF Test			
Mean	Std. Dev	Autocorr.	α	δ	ζ	ρ
3.437	0.206	0.859	0.7529	-0.0009	0.1974	0.7905
(0.033)	(0.024)	(0.077)	(0.3168)	(0.0009)	(0.1587)	(0.0856)

Sample univariate statistics and univariate augmented Dickey-Fuller unit root tests for the log consumption-dividend ratio series. GMM standard errors are in parentheses (1 Newey West lag). For the ADF tests, we estimate the following specification by OLS

$$cd_t = \alpha + \delta t + \zeta \Delta cd_{t-1} + \rho cd_{t-1} + u_t \quad (57)$$

The F-statistic for the joint Wald test, $\delta = 0, \rho = 1$, is 6.16, which is lower than the 5% critical level of 6.25 provided by Dickey and Fuller (1981). The t-statistic for the Wald test, $\rho = 1$, is 6.00 which is higher than the 1% critical value under the null that $\delta = 0$ of 3.96.

All series excluding consumption were obtained from Ibbotson Associates, for 1927-2000, (74 years). Consumption data were obtained from the Bureau of Economic Analysis NIPA tables. Consumption data for the first three years of the sample (1927-1929) are unavailable from the BEA. Aggregate consumption growth was obtained from the website of Robert Shiller, www.econ.yale.edu/~shiller, and used for both nondurables and services consumption series for this period. One observation is lost due to the estimation of models requiring lags. See text and appendix for additional data construction issues.

Table 2: Estimation of the Moody Investor Model

	Means	Feedback	Volatilities	Risk
$E[\Delta d]$	0.044 (0.009)	ρ_d -0.100 (0.058)	σ_d 0.070 (0.016)	$\ln(\beta)$ -0.134 (0.043)
$E[\pi]$	0.034 (0.004)	ρ_u 0.995 (0.014)	σ_u 0.007 (0.006)	γ 1.254 (0.538)
δ	-0.011 (0.009)	ρ_π 0.908 (0.034)	σ_π 0.013 (0.003)	
		ρ_q 0.865 (0.033)	f 1.127 (0.249)	
		ρ_{dq} -0.012 (0.010)	σ_{ud} -0.065 (0.014)	
		ρ_{uq} 0.006 (0.012)	κ_1 0.011 (0.011)	
			κ_2 -0.013 (0.011)	
J-Stat	16.952			
(pval)	(0.075)			

The estimated model is defined by

$$\begin{aligned}
\Delta d_{t+1} &= \mu_d + \rho_d \Delta d_t + \rho_{dq} q_t + \sigma_d \varepsilon_{t+1}^d + \kappa_1 \sqrt{q_t} \varepsilon_{t+1}^q \\
u_{t+1} &= \rho_u u_t + \rho_{uq} q_t + \sigma_{ud} \varepsilon_{t+1}^d + \sigma_u \varepsilon_{t+1}^u + \kappa_2 \sqrt{q_t} \varepsilon_{t+1}^q \\
\pi_{t+1} &= \mu_\pi + \rho_\pi \pi_t + \sigma_\pi \varepsilon_{t+1}^\pi \\
q_{t+1} &= \mu_q + \rho_q q_t + \frac{1}{f} \sqrt{\mu_q (1 - \rho_q^2)} \sqrt{q_t} \varepsilon_{t+1}^q \\
c_{t+1} &= \lambda + \delta t + d_{t+1} + u_{t+1} \implies \Delta c_{t+1} = \delta + \Delta d_{t+1} + \Delta u_{t+1} \\
m_{t+1}^r &= \ln(\beta) - \gamma \Delta c_{t+1} + \gamma \Delta q_{t+1}
\end{aligned}$$

The moments fit are (28 total)

$$\begin{aligned}
&E \left[\Delta d_t, \Delta c_t, \pi_t, r_t^f, dp_t, spd_t, r_t^{ex}, r_t^{bx} \right] \\
&E \left[(\Delta d_t)^2, (\Delta c_t)^2, (\pi_t)^2, (r_t^f)^2, (\Delta d_t \cdot \Delta c_t), (\Delta d_t \cdot r_t^f), (\Delta c_t \cdot r_t^f), (\pi_t \cdot r_t^f) \right] \\
&E \left[(dp_t)^2, (spd_t)^2, (r_t^{ex})^2, (r_t^{bx})^2, (\Delta d_t \cdot r_t^{ex}), (\Delta c_t \cdot r_t^{ex}), (\Delta d_t \cdot r_t^{bx}), (\Delta c_t \cdot r_t^{bx}) \right] \\
&E \left[(\Delta d_t \cdot \Delta d_{t-1}), (\Delta c_t \cdot \Delta c_{t-1}), (\pi_t \cdot \pi_{t-1}), (r_t^f \cdot r_{t-1}^f) \right]
\end{aligned}$$

GMM standard errors are in parentheses (two Newey West lags). The model is estimated in a one-step GMM procedure (see the appendix for details) with a fixed optimal weighting matrix estimated from the data. Data are annual from 1927-2000 (74 years). See data appendix for additional data construction notes.

Table 3: Implied Moments for Moody Investor Model

	Δd_t	Δc_t	π_t	r_t^f	dp_t	spd_t	r_t^{ex}	r_t^{bx}
Mean	[0.045] 0.035 (0.016)	[0.034] 0.032 (0.003)	[0.031] 0.032 (0.006)	[0.051] 0.041 (0.005)	[0.035] 0.039 (0.002)	[0.008] 0.009 (0.002)	[0.060] 0.059 (0.024)	[0.018] 0.010 (0.009)
Std.	[0.072]	[0.013]	[0.031]	[0.030]	[0.012]	[0.013]	[0.176]	[0.085]
Dev.	0.124 (0.022)	0.022 (0.005)	0.042 (0.007)	0.032 (0.004)	0.014 (0.002)	0.013 (0.001)	0.197 (0.022)	0.081 (0.007)
Auto.	[-0.094]	[-0.062]	[0.911]	[0.830]	[0.874]	[0.486]	[-0.049]	[-0.018]
Corr.	0.159 (0.164)	0.414 (0.185)	0.646 (0.160)	0.895 (0.105)	0.800 (0.286)	0.734 (0.106)	0.081 (0.162)	-0.080 (0.164)
Correlations								
Δc_t	[0.49] 0.59 (0.14)							
π_t	[0.01] 0.08 (0.21)	[0.02] 0.43 (0.19)						
r_t^f	[-0.28] -0.01 (0.14)	[-0.16] 0.01 (0.14)	[0.93] 0.42 (0.15)					
dp_t	[-0.04] -0.07 (0.14)	[-0.35] -0.34 (0.15)	[-0.02] -0.10 (0.23)	[0.18] -0.26 (0.11)				
spd_t	[0.60] -0.16 (0.14)	[0.16] -0.17 (0.13)	[-0.72] -0.33 (0.14)	[-0.80] -0.65 (0.10)	[0.30] 0.11 (0.11)			
r_t^{ex}	[0.21] 0.09 (0.13)	[0.21] 0.19 (0.17)	[0.04] 0.01 (0.22)	[-0.06] -0.13 (0.10)	[-0.23] -0.43 (0.14)	[0.03] 0.13 (0.09)		
r_t^{bx}	[-0.03] -0.05 (0.08)	[0.02] -0.06 (0.10)	[-0.36] -0.28 (0.11)	[-0.34] -0.34 (0.13)	[-0.02] -0.09 (0.10)	[0.23] 0.57 (0.07)	[0.39] 0.15 (0.09)	

The numbers in square brackets are simulated moments of the Moody Investor Model. Using the point estimates from Table 2, the system was simulated for 10,000 periods. Dividend yield and excess equity return simulated moments are based upon the log-linear approximation described in the text. The simulated univariate moments of the ‘exact’ dividend yield and excess equity return series are as follows: the mean standard deviation and autocorrelation of the log dividend yield are 0.034, 0.011 and 0.861, and the same statistics for log excess equity returns are 0.060, 0.158 and -0.055 respectively. The second number in each entry is the sample moment based on the annual dataset (1927-2000) and the third number in parentheses is a GMM standard error for the sample moment (one Newey West lag). See data appendix for additional data construction notes.

Table 4: Variance Decomposition Under the Moody Investor Model

	Δd_t	u_t	π_t	q_t	Δd_{t-1}	u_{t-1}	π_{t-1}	q_{t-1}
r_t^f	0.0897	0.0155	0.8649	0.0299	0.0000	0.0000	0.0000	0.0000
dp_t	0.0002	0.0299	0.0000	0.9699	0.0000	0.0000	0.0000	0.0000
spd_t	0.4022	-0.0004	0.5050	0.0932	0.0000	0.0000	0.0000	0.0000
r_t^{ex}	0.0860	-0.0035	0.0055	0.4603	-0.0005	-0.0021	0.0000	0.4543
r_t^{bx}	-0.0046	0.0008	0.7527	0.0346	-0.0001	0.0015	0.0000	0.2152
pd_t	0.0002	0.0299	0.0000	0.9699	0.0000	0.0000	0.0000	0.0000
$E_t \begin{bmatrix} r_{t+1}^{ex} \\ r_{t+1}^{bx} \end{bmatrix}$	0.0000	0.0014	0.0000	0.9986	0.0000	0.0000	0.0000	0.0000
$V_t \begin{bmatrix} r_{t+1}^{ex} \\ r_{t+1}^{bx} \end{bmatrix}$	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000
$E_t \begin{bmatrix} r_{t+1}^{ex} \\ r_{t+1}^{bx} \end{bmatrix}$	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000
$V_t \begin{bmatrix} r_{t+1}^{ex} \\ r_{t+1}^{bx} \end{bmatrix}$	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000
$C_t \begin{bmatrix} r_{t+1}^{ex} \\ r_{t+1}^{bx} \end{bmatrix}$	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000

Under the Moody Investor model, the variable in each row can be expressed as a linear combination of the current and lagged state vector. Generally, for the row variables,

$$x_t = \mu + \Gamma' Y_t^c$$

where Y_t^c is the companion form of Y_t (that is Y_t^c stacks Y_t, Y_{t-1}). Based on μ and Γ , the proportion of the variation of each row variable attributed to the k^{th} element of the state vector is calculated as

$$\frac{\Gamma' v(Y_t^c) \Gamma^{(k)}}{\Gamma' v(Y_t^c) \Gamma}$$

where $\Gamma^{(k)}$ is a column vector with the k^{th} element equal to those of Γ and zero elsewhere. Essentially, the numerator computes the covariance of Y_t with the state variable.

Table 5: Risk Aversion Under the Moody Investor Model

Percentile	1%	5%	10%	25%	50%	75%	90%	95%	99%
	1.11	1.32	1.42	1.79	2.69	5.05	11.3	20.7	91.3
Correlation	Δd_t	Δc_t	π_t	r_t^f	dp_t	spd_t	r_t^{ex}	r_t^{bx}	q_{t-1}
	-0.02	-0.17	-0.03	0.05	0.45	0.16	0.16	-0.11	0.73

Simulated risk aversion moments under the Moody Investor Model. The system was simulated for 10000 periods using the data generating process and point estimates from Table 2. Risk aversion is calculated as

$$RA_t = \gamma \exp(q_t)$$

Table 6A: Correlation Test

Statistic	Sim.	Samp.	p-value
			Cross
$\rho(r_t^{ex}, r_t^{bx})$	[0.391]	0.152 (0.100)	0.021

Table 6B: Predictability Tests (univariate slope coefficients)

Inst.	r_{t+1}^{ex}			r_{t+1}^{bx}			$joint$
	Sim.	p-values		Sim.	p-values		p-values
		Samp.	Cross		Samp.	Cross	
r_t^f	[0.264]	-0.828 (0.639)	0.036	[0.824]	-0.073 (0.399)	0.353	0.089
dp_t	[3.156]	1.691 (2.192)	0.920	[1.455]	0.209 (0.567)	0.333	0.621
spd_t	[0.846]	3.319 (1.726)	0.079	[0.469]	2.120 (0.652)	0.088	0.084
$dp_t - r_t^f$	[0.240]	0.812 (0.643)	0.100	[0.149]	0.079 (0.282)	0.470	0.199

Correlation and predictability performance of the Moody Investor model. The model was simulated for 10000 periods using the point estimates from Table 2. Statistics calculated from the simulated series are reported in square brackets under the heading, ‘Sim.’. For comparison, corresponding sample statistics and GMM standard errors (in parentheses) are reported under the heading ‘Samp.’

The ‘Cross’ column reports GMM based orthogonality test p-values. Recall that both unconditional correlation and single regressor slope coefficients can be expressed as functions of univariate uncentered first and second moments, and a cross moment. Specifically,

$$\rho(x_t^1, x_t^2) = \frac{E[x_t^1 x_t^2] - E[x_t^1] E[x_t^2]}{\left(E[(x_t^1)^2] - (E[x_t^1])^2\right)^{\frac{1}{2}} \left(E[(x_t^2)^2] - (E[x_t^2])^2\right)^{\frac{1}{2}}}$$

$$\beta(x_{t+1}^1, x_t^2) = \frac{E[x_{t+1}^1 x_t^2] - E[x_{t+1}^1] E[x_t^2]}{E[(x_t^2)^2] - (E[x_t^2])^2}$$

In the estimation reported in Table 2, for all the above statistics, the univariate relevant uncentered first and second moments are included as moments to be fit (in the overidentified system), while the cross moments are *not* included in the moment list. Therefore, we focus only on the cross moments for these tests, calculating p-values for each relevant cross-moment. In the case of the slope coefficients, this is $E[x_{t+1}^1 x_t^2]$. In the columns labeled ‘joint,’ slope coefficients for r_t^{ex} and r_t^{bx} are tested jointly. See the appendix for a description of the construction of the GMM based test statistics. Data are annual from 1927-2000, (74 years). See the data appendix for additional data construction notes.

A Data Appendix

In this appendix, we list all the variables used in the article and describe how they were computed from original data sources.

1. r_t^{ex} . To calculate excess equity returns, we start with the CRSP disaggregated monthly stock file, and define monthly US aggregate equity returns as:

$$\begin{aligned} RET_t^m &= \sum_{n=1}^N ret_{i,t} \cdot \frac{(prc_{i,t-1} \cdot shrout_{i,t-1})}{MCAP_{t-1}^m} \\ MCAP_t^m &= \sum_{n=1}^N prc_{i,t} \cdot shrout_{i,t} \end{aligned} \quad (58)$$

where the universe of stocks includes those listed on the AMEX, NASDAQ or NYSE, $ret_{i,t}$ is the monthly total return to equity for a firm, $prc_{i,t}$ is the closing monthly price of the stock, and $shrout_{i,t}$ are the number of shares outstanding at the end of the month for stock i . We create annual end-of-year observations by summing $\ln(1 + RET_t^m)$ over the course of each year. Excess returns are then defined as:

$$r_t^{ex} \equiv \ln(1 + RET_t) - r_{t-1}^f \quad (59)$$

where the risk free rate, r_t^f , is defined below. Note that the lagged risk free rate is applied to match the period over which the two returns are earned (r_t^f is dated when it enters the information set).

2. r_t^{bx} . Excess long bond returns are defined as,

$$r_t^{bx} \equiv \ln(1 + LTBR_t) - r_{t-1}^f \quad (60)$$

where $LTBR_t$ is the annually measured ‘long term government bond holding period return’ from the Ibbottson Associates SBBI yearbook.

3. Δd_t . Log real dividend growth is defined as:

$$\begin{aligned}\Delta d_t &\equiv \ln(DIV_t) - \ln(DIV_{t-1}) - \pi_t \\ DIV_t &= \sum_{month=1}^{12} \left(\sum_{n=1}^N (ret_{i,t} - retx_{i,t}) (prc_{i,t-1} \cdot shrout_{i,t-1}) \right)\end{aligned}\quad (61)$$

where $ret_{i,t}$, $retx_{i,t}$, $prc_{i,t}$ and $shrout_{i,t}$ are total return, total return excluding dividends, price per share and number of common shares outstanding for all issues traded on the AMEX, NASDAQ and AMEX as reported in the CRSP monthly stock files. π_t , inflation, is defined below.

4. spd_t . The yield spread is defined as:

$$spd_t \equiv \ln(1 + LTBY_t) - r_t^f \quad (62)$$

where $LTBY_t$ is the annually measured ‘long term government bond yield’ as reported by Ibbottson Associates in the SBBI yearbook.

5. π_t . Log inflation is defined as:

$$\pi_t \equiv \ln(1 + INFL_t) \quad (63)$$

where $INFL_t$ is the annually measured gross rate of change in the consumer price index as reported by Ibbottson Associates in the SBBI yearbook.

6. *Consumption*. C_t . Total real aggregate consumption is calculated as total constant dollar non-durable plus services consumption as reported in the NIPA tables available from the website of the US Bureau of Economic Analysis. As described in Section 2, we checked the robustness of our results to the use of alternative consumption measures that more closely approximate the consumption of stockholders. C_t^{LX} denotes real luxury consumption, defined as the sum of three disaggregated constant-dollar NIPA consumption series: boats and aircraft, (C_t^{BA}), jewelry and watches, (C_t^{JW}) and foreign travel, (C_t^{FT}). C_t^{WT} , ‘participation weighted consumption,’ is defined as follows

$$C_t^{WT} = PART_t \cdot (C_t^{AG} - C_t^{FT}) + C_t^{LX} \quad (64)$$

The series should more accurately reflect the consumption basket of stock market participants. The higher the stock market participation rate, $PART_t$, the more relevant is aggregate (non-luxury) consumption. C_t^{FT} is subtracted from C_t^{AG} to avoid double-counting (the other elements of luxury consumption are classified as durables, and thus not included in total nondurable and service consumptions, which comprise C_t^{AG}). $PART_t$ is the US stock market participation rate taken from data provided by Steve Zeldes (see Ameriks and Zeldes (2002)): the percent of US households with direct or indirect ownership of stocks:

Year	Rate
1962	0.296
1983	0.437
1989	0.475
1992	0.496
1998	0.570

(65)

From these data, an interpolated participation rate, $PART_t$ was calculated by the authors as $(1 + \exp[-(-62.0531 + 0.03118 \cdot YEAR_t)])^{-1}$, the result of estimating a deterministic trend line through the numbers in the table above. To fill in consumption data prior to NIPA coverage, 1926-1928 (inclusive), we applied (in real terms) the growth rate of real consumption reported at the website of Robert Shiller for those years. The real log growth rates of all consumption series are calculated as:

$$\Delta c_t = \ln(C_t) - \ln(C_{t-1}) - \pi_t \quad (66)$$

Note that the same inflation series, defined above, is applied to deflate all three consumption measures.

7. dp_t . The dividend yield measure used in this paper is:

$$\begin{aligned} dp_t &\equiv \ln\left(1 + \frac{DIV_t}{MCAP_t}\right) \\ MCAP_t &= MCAP_{t,DEC}^m \end{aligned} \quad (67)$$

where DIV_t is defined above and $MCAP_{t,DEC}^m$ corresponds to the December value of $MCAP_t^m$ for each year.

8. r_t^f . The short term risk free rate is defined as:

$$r_t^f \equiv \ln(1 + STBY_t) \quad (68)$$

where $STBY_t$ is the ‘short term government bond yield’ reported by the St. Louis federal reserve statistical release website (FRED). From this monthly series, we took December values to create annual end of year observations. Note that r_t^f is dated when it enters the information set, the end of the period prior to that over which the return is earned. For instance, the risk free rate earned from January 1979 through December 1979 is dated as (end-of-year) 1978.

B The General Pricing Model

Here we collect proofs of all the pricing propositions. For completeness, we report the general pricing model equations. 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.

B.1 Definition of the System

The state vector is described by,

$$\begin{aligned} Y_t &= \mu + AY_{t-1} + (\Sigma_F F_{t-1} + \Sigma_H) \varepsilon_t \\ F_t &= (\parallel \phi + \Phi Y_t \parallel)' \odot I, \end{aligned} \tag{69}$$

where Y_t is the state vector of length k , μ and ϕ are parameter vectors also of length k and A , Σ_F , Σ_H and Φ are parameter matrices of size $(k \times k)$. ε_t is a k -vector of zero mean i.i.n. innovations. The log of the real stochastic discount factor is modeled as,

$$m_{t+1}^r = \mu_m + \Gamma_m' Y_t + (\Sigma_{mF}' F_t + \Sigma_{mH}') \varepsilon_{t+1}$$

where μ_m , is scalar and Γ_m , Σ_{mF} , and Σ_{mH} are k-vectors of parameters. The following restrictions are required:

$$\begin{aligned}
\Sigma_F F_t \Sigma_H' &= 0 \\
\Sigma_{mF}' F_t \Sigma_{mH} &= 0 \\
\Sigma_H F_t \Sigma_{mF} &= 0 \\
\Sigma_F F_t \Sigma_{mH} &= 0 \\
\phi + \Phi Y_t &\geq 0
\end{aligned} \tag{70}$$

These restrictions are convenient for the calculation of conditional expectations of functions of Y_t .

B.2 Some Useful Lemmas

Lemma 1. The conditional expectation of an exponential affine function of the state variables and the pricing kernel is given by:

$$E_t [\exp (a + c' Y_{t+1} + m_{t+1}^r)] = \exp (g_0 + g' Y_t) \tag{71}$$

Proof. By lognormality,

$$E_t [\exp (a + c' Y_{t+1} + m_{t+1}^r)] = \exp \left(\begin{aligned} &a + c' E_t [Y_{t+1}] + E_t [m_{t+1}^r] \\ &+ \frac{1}{2} (V_t [c' Y_{t+1}] + V_t [m_{t+1}^r] + 2C_t [c' Y_{t+1}, m_{t+1}^r]) \end{aligned} \right) \tag{72}$$

We will take each of the five conditional expectations above separately. Below, \odot (the Hadamard product) denotes element-by-element multiplication.

1.

$$c' E_t [Y_{t+1}] = c' (\mu + A Y_t) \tag{73}$$

2.

$$E_t [m_{t+1}^r] = \mu_m + \Gamma_m' Y_t \tag{74}$$

3.

$$\begin{aligned}
V_t [c'Y_{t+1}] &= c' (\Sigma_F F_t + \Sigma_H) (\Sigma_F F_t + \Sigma_H)' c \\
&= c' \Sigma_F F_t F_t' \Sigma_F' c + c' \Sigma_H \Sigma_H' c \\
&= ((\Sigma_F' c) \odot (\Sigma_F' c))' (\phi + \Phi Y_t) + c' \Sigma_H \Sigma_H' c
\end{aligned} \tag{75}$$

where the second line uses restrictions in Equation (70) and the third line follows from properties of the \odot operator.

4.

$$\begin{aligned}
V_t [m_{t+1}^r] &= (\Sigma_{mF}' F_t + \Sigma_{mH}') (\Sigma_{mF}' F_t + \Sigma_{mH}')' \\
&= \Sigma_{mF}' F_t F_t' \Sigma_{mF} + \Sigma_{mH}' \Sigma_{mH} \\
&= (\Sigma_{mF} \odot \Sigma_{mF})' (\phi + \Phi Y_t) + \Sigma_{mH}' \Sigma_{mH}
\end{aligned} \tag{76}$$

where the second line uses restrictions in Equation (70) and the third line follows from properties of the \odot operator.

5.

$$\begin{aligned}
C_t [c'Y_{t+1}, m_{t+1}^r] &= c' [(\Sigma_F F_{t-1} + \Sigma_H) (\Sigma_{mF}' F_t + \Sigma_{mH}')'] \\
&= c' [\Sigma_F F_t F_t' \Sigma_{mF} + \Sigma_H \Sigma_{mH}] \\
&= c' [(\Sigma_{mF}' \odot \Sigma_F) (\phi + \Phi Y_t) + \Sigma_H \Sigma_{mH}']
\end{aligned} \tag{77}$$

where the second line uses restrictions in Equation (70) and the third line follows from properties of the \odot operator.

Substituting,

$$\begin{aligned}
& E_t [\exp (a + c' Y_{t+1} + m_{t+1}^r)] \\
&= \exp \left(\begin{aligned} & a + c' (\mu + A Y_t) + \mu_m + \Gamma'_m Y_t \\ & + \frac{1}{2} ((\Sigma'_F c) \odot (\Sigma'_F c))' (\phi + \Phi Y_t) + \frac{1}{2} c' \Sigma_H \Sigma'_H c \\ & + \frac{1}{2} (\Sigma_{mF} \odot \Sigma_{mF})' (\phi + \Phi Y_t) + \frac{1}{2} \Sigma_{mH} \Sigma'_{mH} + \frac{1}{2} \sigma_m^2 \\ & + c' [(\Sigma'_{mF} \odot \Sigma_F) (\phi + \Phi Y_t) + \Sigma_H \Sigma'_{mH}] \end{aligned} \right) \quad (78)
\end{aligned}$$

Evidently,

$$\begin{aligned}
g_0 &= a + c' \mu + \mu_m + \frac{1}{2} ((\Sigma'_F c) \odot (\Sigma'_F c))' \phi \\
&\quad + \frac{1}{2} c' \Sigma_H \Sigma'_H c + \frac{1}{2} (\Sigma_{mF} \odot \Sigma_{mF})' \phi + \frac{1}{2} \Sigma'_{mH} \Sigma_{mH} + \frac{1}{2} \sigma_m^2 + c' [(\Sigma'_{mF} \odot \Sigma_F) \phi + \Sigma_H \Sigma_{mH}] \\
g' &= c' A + \Gamma'_m + \frac{1}{2} ((\Sigma'_F c) \odot (\Sigma'_F c))' \Phi + \frac{1}{2} (\Sigma_{mF} \odot \Sigma_{mF})' \Phi + c' [(\Sigma'_{mF} \odot \Sigma_F) \Phi] \quad (79)
\end{aligned}$$

B.3 Representation of the Nominal Risk Free Rate

It is well known that the gross risk free rate is given by the inverse of the conditional expectation of the nominal pricing kernel. Let e_π denote the k-vector which selects log inflation from the state vector. Then,

$$\exp(r_t^f) = (E_t [\exp(m_{t+1}^r - e'_\pi Y_{t+1})])^{-1} \quad (80)$$

Applying Lemma 1 with $a = 0$ and $c = -e_\pi$, it is immediate that

$$r_t^f = -a_1^0 - a_1' Y_t \quad (81)$$

where a_0 and a_1' are given by the Lemma 1.

B.4 Representation of the Entire Term Structure

The proof to demonstrate the affine form for the term structure is accomplished by induction.

Recall that the nominal one period risk free rate is given by,

$$\begin{aligned} r_t^f &= -a_1^0 - a_1' Y_t \\ p_{1,t} &= a_1^0 + a_1' Y_t \end{aligned} \tag{82}$$

where a_1^0 and a_1' are given in the previous subsection. Recall also the recursive relation of discount bond prices in the stochastic discount factor representation,

$$P_{n,t} = E_t [M_{t+1}^N P_{n-1,t+1}] . \tag{83}$$

Suppose, for the purposes of induction, that $P_{n-1,t}$ can be expressed as,

$$P_{n-1,t} = \exp(a_{n-1}^0 + a_{n-1}' Y_t) . \tag{84}$$

Then, leading this expression by one period and substituting it into the recursive relation, (83), we have,

$$\begin{aligned} P_{n,t} &= E_t [\exp(m_{t+1}^N + p_{n-1,t+1})] \\ &= E_t [\exp(m_{t+1}^r - e_\pi' Y_{t+1} + a_{n-1}^0 + a_{n-1}' Y_{t+1})] \\ &= \exp(a_n^0 + a_n' Y_t) \end{aligned} \tag{85}$$

where the coefficients are given (recursively) by Lemma 1. Upon substitution, the following recursion is revealed

$$\begin{aligned} a_n^0 &= a_{n-1}^0 + (a_{n-1} - e_\pi)' \mu + \mu_m + \frac{1}{2} ((\Sigma_F' (a_{n-1} - e_\pi)) \odot (\Sigma_F' (a_{n-1} - e_\pi)))' \phi \\ &\quad + \frac{1}{2} (a_{n-1} - e_\pi)' \Sigma_H \Sigma_H' (a_{n-1} - e_\pi) + \frac{1}{2} (\Sigma_{mF} \odot \Sigma_{mF})' \phi \\ &\quad + \frac{1}{2} \Sigma_{mH}' \Sigma_{mH} + \frac{1}{2} \sigma_m^2 + (a_{n-1} - e_\pi)' [(\Sigma_{mF}' \odot \Sigma_F) \phi + \Sigma_H \Sigma_{mH}] \\ a_n' &= (a_{n-1} - e_\pi) A + \Gamma_m' + \frac{1}{2} ((\Sigma_F' (a_{n-1} - e_\pi)) \odot (\Sigma_F' (a_{n-1} - e_\pi)))' \Phi \\ &\quad + \frac{1}{2} (\Sigma_{mF} \odot \Sigma_{mF})' \Phi + (a_{n-1} - e_\pi) [(\Sigma_{mF}' \odot \Sigma_F) \Phi] \end{aligned} \tag{86}$$

B.5 Representation of the Equity Prices

To demonstrate the dependence of the price-dividend ratio on Y_t , we use a proof by induction. Let e_d be the k-vector which selects real dividend growth from Y_t . The price dividend ratio is given by

$$\begin{aligned}
\frac{P_t}{D_t} &= E_t \sum_{n=1}^{\infty} \exp \left(\sum_{j=1}^n m_{t+j}^r + g d_{t+j} \right) \\
&= E_t \sum_{n=1}^{\infty} \exp \left(\sum_{j=1}^n m_{t+j}^r + e_d' Y_{t+j} \right) \\
&= \sum_{n=1}^{\infty} q_{n,t}^0
\end{aligned} \tag{87}$$

where $q_{n,t}^0 \equiv E_t \exp \left(\sum_{j=1}^n m_{t+j}^r + e_d' Y_{t+j} \right)$ are scalars. We will prove that $q_{n,t}^0 = \exp (b_n^0 + b_n' Y_t)$ where b_n^0 (scalar) and b_n' (k-vectors) are defined below. The proof is accomplished by induction. Consider $q_{1,t}^0$:

$$q_{1,t}^0 = E_t \left(\exp (m_{t+1}^r + e_d' Y_{t+1}) \right) \tag{88}$$

By Lemma 1,

$$q_{1,t}^0 = \exp (b_1^0 + b_1' Y_t) \tag{89}$$

where b_1^0 and b_1' are given by Lemma 1. Next, suppose that $q_{n-1,t}^0 = \exp (b_{n-1}^0 + b_{n-1}' Y_t)$. Then rearrange $q_{n,t}^0$ as follows.

$$\begin{aligned}
q_{n,t}^0 &= E_t \exp \left(\sum_{j=1}^n m_{t+j}^r + e_d' Y_{t+j} \right) \\
&= E_t E_{t+1} \left\{ \exp (m_{t+1}^r + e_d' Y_{t+1}) \exp \left(\sum_{j=1}^{n-1} m_{t+j+1}^r + e_d' Y_{t+j+1} \right) \right\} \\
&= E_t \left\{ \exp (m_{t+1}^r + e_d' Y_{t+1}) E_{t+1} \exp \left(\sum_{j=1}^{n-1} m_{t+j+1}^r + e_d' Y_{t+j+1} \right) \right\} \\
&= E_t \left\{ \exp (m_{t+1}^r + e_d' Y_{t+1}) q_{n-1,t+1}^0 \right\} \\
&= E_t \left\{ \exp (m_{t+1}^r + e_d' Y_{t+1}) \exp (b_{n-1}^0 + b_{n-1}' Y_{t+1}) \right\} \\
&= E_t \left\{ \exp (b_{n-1}^0 + m_{t+1}^r + (e_d + b_{n-1})' Y_{t+1}) \right\} \\
&= \exp (b_n^0 + b_n' Y_t)
\end{aligned} \tag{90}$$

where b_n^0 and b'_n are easily calculated using Lemma 1. Upon substitution, the recursions are revealed to be given by,

$$\begin{aligned}
b_n^0 &= b_{n-1}^0 + (b_{n-1} + e_d)' \mu + \mu_m + \frac{1}{2} ((\Sigma'_F (b_{n-1} + e_d)) \odot (\Sigma'_F (b_{n-1} + e_d)))' \phi \\
&\quad + \frac{1}{2} (b_{n-1} + e_d)' \Sigma_H \Sigma'_H (b_{n-1} + e_d) + \frac{1}{2} (\Sigma_{mF} \odot \Sigma_{mF})' \phi \\
&\quad + \frac{1}{2} \Sigma'_{mH} \Sigma_{mH} + \frac{1}{2} \sigma_m^2 + (b_{n-1} + e_d)' [(\Sigma'_{mF} \odot \Sigma_F) \phi + \Sigma_H \Sigma_{mH}] \\
b'_n &= (b_{n-1} + e_d)' A + \Gamma'_m + \frac{1}{2} ((\Sigma'_F (b_{n-1} + e_d)) \odot (\Sigma'_F (b_{n-1} + e_d)))' \Phi \\
&\quad + \frac{1}{2} (\Sigma_{mF} \odot \Sigma_{mF})' \Phi + (b_{n-1} + e_d) [(\Sigma'_{mF} \odot \Sigma_F) \Phi]
\end{aligned} \tag{91}$$

For the purposes of estimation the coefficient sequences are calculated out 200 years. If the resulting calculated value for PD_t has not converged, then the sequences are extended another 100 years until either the PD_t value converges, or becomes greater than 1000 in magnitude.

C The Estimated Model

The state variables and dynamics for the estimated model are given by

$$\begin{aligned}
q_{t+1} &= \mu_q + \rho_q q_t + \frac{1}{f} \sqrt{\mu_q (1 - \rho_q^2)} \sqrt{q_t} \varepsilon_{t+1}^q \\
\Delta d_{t+1} &= \mu_d + \rho_d \Delta d_t + \rho_{dq} q_t + \sigma_d \varepsilon_{t+1}^d + \kappa_1 \sqrt{q_t} \varepsilon_{t+1}^q \\
u_{t+1} &= 0 + \rho_u u_t + \rho_{uq} q_t + \sigma_{ud} \varepsilon_{t+1}^d + \sigma_u \varepsilon_{t+1}^u + \kappa_2 \sqrt{q_t} \varepsilon_{t+1}^q \\
\Delta c_{t+1} &= \delta + \Delta d_{t+1} + \Delta u_{t+1} \\
\pi_{t+1} &= \mu_\pi + \rho_\pi \pi_t + \sigma_\pi \varepsilon_{t+1}^\pi
\end{aligned} \tag{92}$$

$$m_{t+1} = \ln(\beta) - \gamma \cdot \Delta c_{t+1} + \gamma \cdot (q_{t+1} - q_t) \tag{93}$$

This model is clearly a special case of the model above. The implied system matrices are,

$$\mu = \begin{bmatrix} \mu_q \\ \mu_d \\ 0 \\ \mu_\pi \end{bmatrix} \quad A = \begin{bmatrix} \rho_q & 0 & 0 & 0 \\ \rho_{dq} & \rho_d & 0 & 0 \\ \rho_{uq} & 0 & \rho_u & 0 \\ 0 & 0 & 0 & \rho_\pi \end{bmatrix} \quad \Sigma_H = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \sigma_d & 0 & 0 \\ 0 & \sigma_{ud} & \sigma_u & 0 \\ 0 & 0 & 0 & \sigma_\pi \end{bmatrix} \quad \Sigma_F = \begin{bmatrix} \sigma_q & 0 & 0 & 0 \\ \kappa_1 & 0 & 0 & 0 \\ \kappa_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\phi = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \Phi = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \tag{94}$$

where $\sigma_q \equiv \frac{1}{f} \sqrt{\mu_q (1 - \rho_q^2)}$. Note that in this formulation, f is the ratio of the unconditional mean to the unconditional variance of q_t . It is estimated directly and reported.

The pricing kernel for the model can be written as:

$$\begin{aligned}
m_{t+1} &= \ln(\beta) - \gamma(\delta + \Delta d_{t+1} + \Delta u_{t+1}) + \gamma(q_{t+1} - q_t) \\
&= \ln(\beta) - \gamma \begin{pmatrix} -(\delta + \mu_d + \rho_d \Delta d_t + \rho_{dq} q_t + \sigma_d \varepsilon_{t+1}^d + \kappa_1 \sqrt{q_t} \varepsilon_{t+1}^q) + \dots \\ -(\rho_u u_t + \rho_{uq} q_t + \sigma_{ud} \varepsilon_{t+1}^d + \sigma_u \varepsilon_{t+1}^u + \kappa_2 \sqrt{q_t} \varepsilon_{t+1}^q) + \dots \\ \mu_q + \rho_q q_t + \frac{1}{f} \sqrt{\mu_q (1 - \rho_q^2)} \sqrt{q_t} \varepsilon_{t+1}^q - q_t - u_t \end{pmatrix}
\end{aligned} \tag{95}$$

We can now read off the pricing kernel matrices:

$$\begin{aligned}
\mu_m &= (\ln(\beta) - \gamma\delta - \gamma\mu_d + \gamma\mu_q) & \Gamma_m &= \gamma \begin{bmatrix} -\rho_{dq} - \rho_{uq} + \rho_q - 1 \\ -\rho_d \\ 1 - \rho_u \\ 0 \end{bmatrix} & \Sigma_{mH} &= \gamma \begin{bmatrix} 0 \\ -\sigma_d - \sigma_{ud} \\ -\sigma_u \\ 0 \end{bmatrix} \\
\Sigma_{mF} &= \gamma \begin{bmatrix} -\kappa_1 - \kappa_2 + \sigma_q \\ 0 \\ 0 \\ 0 \end{bmatrix}.
\end{aligned} \tag{96}$$

C.1 Alternate models

Several other models were explored during this study. First, the alternate consumption measures Δc_t^{lx} and Δc_t^{wt} were tried in place of aggregate consumption. The results for weighted consumption were nearly identical to those reported above. For luxury consumption, the model failed to converge. This may be due to a lack of cointegration between aggregate dividends and the small component of consumption represented by the few luxury series we identified.

Secondly, models wherein consumption growth and dividend growth are not cointegrated were considered. Specifically, attempts were made to estimate models of the form

$$\begin{aligned}
\Delta d_{t+1} &= \mu_d + \rho_d \Delta d_t + \rho_{dq} q_t + \sigma_d \varepsilon_{t+1}^d + \kappa_1 \sqrt{q_t} \varepsilon_{t+1}^q \\
\Delta c_{t+1} &= \mu_c + \rho_c \Delta c_t + \rho_{cq} q_t + \sigma_{cd} \varepsilon_{t+1}^d + \sigma_c \varepsilon_{t+1}^c + \kappa_2 \sqrt{q_t} \varepsilon_{t+1}^q \\
\pi_{t+1} &= \mu_\pi + \rho_\pi \pi_t + \sigma_\pi \varepsilon_{t+1}^\pi \\
q_{t+1} &= \mu_q + \rho_q q_t + \frac{1}{f} \sqrt{\mu_q (1 - \rho_q^2)} \sqrt{q_t} \varepsilon_{t+1}^q \\
m_{t+1}^r &= \ln(\beta) - \gamma \Delta c_{t+1} + \gamma \Delta q_{t+1}
\end{aligned} \tag{97}$$

Estimation was attempted for each of the three consumption measures. However, none of these models converged. This is almost certainly due to the very different pricing implications of a non-stationary consumption-dividend ratio.

C.2 Log Linear Approximation of Equity Prices

In the estimation, we use a linear approximation to the price-dividend ratio. From Equation (87), we see that the price dividend ratio is given by

$$\begin{aligned}\frac{P_t}{D_t} &= \sum_{n=1}^{\infty} q_{n,t}^0 \\ &= \sum_{n=1}^{\infty} \exp(b_n^0 + b'_n Y_t)\end{aligned}\tag{98}$$

and the coefficient sequences, $\{b_n^0\}_{n=1}^{\infty}$ and $\{b'_n\}_{n=1}^{\infty}$, are given above. We seek to approximate the log price-dividend ratio using a first order Taylor approximation of Y_t about \bar{Y} , the unconditional mean of Y_t . Let

$$\bar{q}_n^0 = \exp(b_n^0 + b'_n \bar{Y})\tag{99}$$

and note that

$$\frac{\partial}{\partial Y_t} \left(\sum_{n=1}^{\infty} q_{n,t}^0 \right) = \sum_{n=1}^{\infty} \frac{\partial}{\partial Y_t} q_{n,t}^0 = \sum_{n=1}^{\infty} q_{n,t}^0 \cdot b'_n\tag{100}$$

Approximating,

$$\begin{aligned}pd_t &\simeq \ln \left(\sum_{n=1}^{\infty} \bar{q}_n^0 \right) + \frac{1}{\sum_{n=1}^{\infty} \bar{q}_n^0} \left(\sum_{n=1}^{\infty} \bar{q}_n^0 \cdot b'_n \right) (Y_t - \bar{Y}) \\ &= d_0 + d' Y_t\end{aligned}\tag{101}$$

where d_0 and d' are implicitly defined. Similarly,

$$\begin{aligned}gpd_t &\equiv \ln \left(1 + \frac{P_t}{D_t} \right) \simeq \ln \left(1 + \sum_{n=1}^{\infty} \bar{q}_n^0 \right) + \frac{1}{1 + \sum_{n=1}^{\infty} \bar{q}_n^0} \left(\sum_{n=1}^{\infty} \bar{q}_n^0 \cdot b'_n \right) (Y_t - \bar{Y}) \\ &= h_0 + h' Y_t\end{aligned}\tag{102}$$

where h_0 and h' are implicitly defined. Note also that the dividend yield measure used in this study can be expressed as follows

$$dp_t \equiv \ln \left(1 + \frac{D_t}{P_t} \right) = gpd_t - pd_t\tag{103}$$

so that it is also linear in the state vector under these approximations. Also, log excess equity returns can be represented follows. Using the definition of excess equity returns,

$$\begin{aligned}
r_{t+1}^x &= -r_t^f - pd_t + gd_{t+1} + \pi_{t+1} + gpd_{t+1} \\
&\sim (h_0 - d_0) + (e'_d + e'_\pi + h') Y_{t+1} + (-e'_{rf} + -d') Y_t \\
&= r_0 + r'_1 Y_{t+1} + r'_2 Y_t
\end{aligned} \tag{104}$$

where r_0 , r'_1 and r'_2 are implicitly defined.

C.3 Accuracy of the Price Dividend Ratio Approximation

To assess the accuracy of the log linear approximation of the price dividend ratio, the following experiment was conducted. For the model and point estimates reported in Table 2, a simulation was run for 10,000 periods. In each period, the ‘exact’ price dividend ratio and log dividend yield are calculated in addition to their approximate counterparts derived in the previous subsection. The resulting series for exact and approximate dividend yields and excess stock returns compare as follows:

	appx dp_t	exact dp_t	appx r_t^x	exact r_t^x
mean	0.0355	0.0352	0.0602	0.0599
std. dev.	0.0126	0.0127	0.1579	0.1768
correlation	0.9996		0.9583	

C.4 Analytic Moments of Y_t and W_t

Recall that the data generating process for Y_t is given by,

$$\begin{aligned}
Y_t &= \mu + AY_{t-1} + (\Sigma_F F_{t-1} + \Sigma_H) \varepsilon_t \\
F_t &= \text{sqrt}(\text{diag}(\phi + \Phi Y_t))
\end{aligned} \tag{105}$$

It is straightforward to show that the uncentered first, second, and first autocovariance moments of Y_t are given by,

$$\begin{aligned}\overline{Y_t} &= (I_k - A)^{-1} \mu \\ \text{vec}(\overline{Y_t Y_t'}) &= (I_{k^2} - A \otimes A)^{-1} \cdot \text{vec}(\mu \mu' + \mu \overline{Y_t'} A' + A \overline{Y_t} \mu' + \Sigma_F \overline{F_t^2} \Sigma_F' + \Sigma_H \Sigma_H') \\ \text{vec}(\overline{Y_t Y_{t-1}'}) &= (I_{k^2} - A \otimes A)^{-1} \cdot \text{vec}(\mu \mu' + \mu \overline{Y_t'} A' + A \overline{Y_t} \mu' + A (\Sigma_F \overline{F_t^2} \Sigma_F' + \Sigma_H \Sigma_H'))\end{aligned}\quad (106)$$

where overbars denote unconditional means and $\overline{F_t^2} = \text{diag}(\phi + \Phi \overline{Y_t})$.

Now consider the unconditional moments of a n-vector of observable variables W_t which obey the condition

$$W_t = \mu^w + \Gamma^w Y_{t-1} + (\Sigma_F^w F_{t-1} + \Sigma_H^w) \varepsilon_t \quad (107)$$

where μ^w is an n-vector and Σ_F^w , Σ_H^w and Γ^w are $(n \times k)$ matrices. It is straightforward to show that the uncentered first, second, and first autocovariance moments of W_t are given by,

$$\begin{aligned}\overline{W_t} &= \mu^w + \Gamma^w \overline{Y_t} \\ \overline{W_t W_t'} &= \mu^w \mu^{w'} + \mu^w \overline{Y_t'} \Gamma^{w'} + \Gamma^w \overline{Y_t} \mu^{w'} + \Gamma^w \overline{Y_t Y_t'} \Gamma^{w'} + \Sigma_F^w \overline{F_t^2} \Sigma_F^{w'} + \Sigma_H^w \Sigma_H^{w'} \\ \overline{W_t W_{t-1}'} &= \mu^w \mu^{w'} + \mu^w \overline{Y_t'} \Gamma^{w'} + \Gamma^w \overline{Y_t} \mu^{w'} + \Gamma^w \overline{Y_t Y_{t-1}'} \Gamma^{w'} + \Gamma^w (\Sigma_F^w \overline{F_t^2} \Sigma_F^{w'} + \Sigma_H^w \Sigma_H^{w'})\end{aligned}\quad (108)$$

It remains to demonstrate that the observable series used in estimation obey Equation (107). This is trivially true for elements of W_t which are also elements of Y_t such as Δd_t , Δc_t , π_t . Using Equations (81), (104) and (101), it is apparent that r_t^f , dp_t and r_t^x satisfy Equation (107) as well.