Stock and Bond Returns with Moody Investors

Geert Bekaert*
Columbia University and NBER

Eric Engstrom
Federal Reserve Board of Governors

Steven R. Grenadier
Stanford University and NBER

This Draft: March 2010

JEL Classifications
G12, G15, E44

Keyphrases
Equity premium, Excess Volatility, Stock-Bond Return Correlation, Return Predictability, Countercyclical risk aversion, Habit persistence,

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 the general method of moments, 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, John Campbell, and participants at the NBER Summer Institute 2004 for useful comments. All remaining errors are the sole responsibility of the authors. This work does not necessarily reflect the views of the Federal Reserve System or its staff.

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 counter-cyclical 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 (GMM), we find that our model can rather successfully fit many features of bond and stock return data together with important properties of the fundamentals, including a low correlation between fundamentals and returns.

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. Wachter (2006) and Buraschi and Jiltsov (2007) do provide extensions of the Campbell-Cochrane framework but focus almost entirely on term structure puzzles. In our model, stochastic risk aversion is not perfectly negatively correlated with consumption growth as in Campbell and Cochrane, but the perfect correlation case represents a testable restriction of our model.

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.

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 (2006) 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, Piazessi, Schneider and Tuzel (2007) and Lustig and Van Nieuwerburgh (2005) 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.

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:

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

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 ratio[3]. As the surplus ratio goes to zero, the consumer's risk aversion goes to infinity. In our model, we view the inverse of the surplus ratio as a preference shock, which we denote by Q_t. Thus, $Q_t = \frac{C_t}{C_t-H_t}$. Risk aversion is now characterized by $\gamma \cdot Q_t$, and $Q_t > 1$. 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:

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

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 because it can generate counter-cyclical expected returns and prices of risk. To see this, first

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

[3] Of course, this is not actual 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 scalor 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."

note that the coefficient of variation of the pricing kernel equals the maximum Sharpe ratio attainable with the available assets (see Hansen and Jagannathan, 1991). As Campbell and Cochrane (1999) also note, with a log-normal kernel:

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

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.

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 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,

$$q_{t+1} = \mu_q + \rho_{qq} q_t + \sigma_{qq}\sqrt{q_t}\left(\left(1-\lambda^2\right)^{1/2}\varepsilon_{t+1}^q + \lambda\varepsilon_{t+1}^c\right), \tag{4}$$

where μ_q, ρ_q and σ_q and λ are parameters[4]. Here, ε_t^q is a standard normal innovation process specific to q_t and ε_t^c is a similar process, representing the sole source of conditional uncertainty in the consumption growth process. Both are distributed as $N(0,1)$. We will shortly see that $\lambda \in [-1,1]$ is the conditional correlation between consumption growth and q_t. When $\lambda = -1$, q_t and consumption growth will be perfectly negatively correlated which is consistent with the habit persistence formulation of Campbell and Cochrane (1999). 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.

[4] σ_q is parameterized as $\frac{1}{f}\sqrt{\overline{q}\left(1-\rho_{qq}^2\right)}$ where $\overline{q} = \left(1-\rho_q\right)^{-1}\mu_q$. 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).

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 or consumption process become critical. Because we price equities in this article, dividend growth must be a state variable. Section 2.2.1 details the modeling of consumption and dividend growth. 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. Realistically, consumption is financed by many sources of income (especially labor income) not represented in aggregate dividends.[5]. We therefore represent dividends as consumption divided by the consumption-dividend ratio CD_t. Because dividends and consumption are non-stationary we model consumption growth and the consumption-dividend ratio, CD_t. The main econometric issue is whether CD_t is stationary or, more generally, whether consumption and dividends are cointegrated. Bansal, Dittmar and Lundblad (2002) recently argue that dividends and consumption are cointegrated, but with a cointegrating vector that differs from $[1, -1]$, whereas Bansal and Yaron (2004) assume two unit roots. 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.

We use aggregate nondurables and services consumption as the consumption measure. Because many agents in the economy do not hold stocks at all, we checked the robustness of the model to an alternative measure of consumption that attempts to approximate the consumption of the stockholder. Mankiw and Zeldes (1991) and Ait-Sahalia, Parker and Yogo (2004) have pointed

[5] In the NBER version of this article, we provide a more formal motivation for our set-up in the context of a multiple dividend economy. Menzly, Santos and Veronesi (2004) formulate a continuous-time economy extending the Campbell-Cochrane framework to multiple dividend processes.

out that aggregate consumption may not be representative of the consumption of stock holders. 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 (2004). However, our model does not perform noticeably better with this consumption measure and we do not report these results to conserve space.

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

$$\begin{aligned} \Delta c_{t+1} &= \mu_c + \rho_{cc}\Delta c_t + \rho_{cu} u_t + \sigma_{cc}\sqrt{q_t}\varepsilon^c_{t+1} \\ cd_{t+1} &= \zeta + \delta t + u_{t+1} \end{aligned} \tag{5}$$

where μ_c, ρ_{cc}, ρ_{cu}, and σ_{cc} are parameters governing consumption growth, Δc_t. This specification implies that consumption growth is an ARMA(1,1) processes. Bansal and Yaron (2004) have recently stressed the importance of allowing an MA component in the dividend process, and Wachter (2006) also models consumption growth as an ARMA(1,1) process. Note that we have allowed for heteroskedasticity in the consumption process as the conditional volatility of Δc_{t+1} is proportional to q_t. While this is primarily for modelling convenience in arriving at closed form solutions for asset prices, there is substantial evidence for such heteroskedasticity even in annual real consumption growth, which has (unreported) unconditional excess kurtosis of 7.00 in our sample. During estimation, we are careful to check that our model implied consumption growth kurtosis does not exceed that in the data. The constant ζ 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,

$$\Delta d_{t+1} = \Delta c_{t+1} - \delta - \Delta u_{t+1} \tag{6}$$

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

$$u_{t+1} = \mu_u + \rho_{uu} u_t + \rho_{uc}\Delta c_t + \sigma_{uc}\sqrt{q_t}\varepsilon^c_{t+1} + \sigma_{uu}\varepsilon^u_{t+1}. \tag{7}$$

The conditional covariance between consumption growth and preference shocks can now be more explicitly examined. In particular, this covariance equals:

$$Cov_t\left[\Delta c_{t+1}, q_{t+1}\right] = \sigma_{qq}\sigma_{cc}\lambda q_t \tag{8}$$

so that the covariance is most negative when $\lambda = -1$, a restriction of perfectly counter-cyclical risk aversion under which our model most closely approaches that of Campbell and Cochrane (1999). 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 $\left(\frac{\sigma_{qq}}{\sigma_{cc}}\right)\lambda > 1 - Q_t$ for all t. Note that the right-hand side is negative and 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:

$$\pi_{t+1} = \mu_\pi + \rho_\pi \pi_t + \sigma_\pi \varepsilon^\pi_{t+1} \tag{9}$$

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). These assumptions impose that the Fisher Hypothesis holds in our economy. The pricing of nominal assets then occurs with a nominal pricing kernel, $\widehat{m}_{t+1}$ that is a simple transformation of the real pricing kernel, m_{t+1}.

$$\widehat{m}_{t+1} = m_{t+1} - \pi_{t+1} \tag{10}$$

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:

$$m_{t+1} = \ln(\beta) - \gamma\Delta c_{t+1} + \gamma\Delta q_{t+1} \tag{11}$$

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

$$1 = E_t\left[M_{t+1}\left(1 + R_{i,t+1}\right)\right] \tag{12}$$

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:

$$\begin{aligned}
q_{t+1} &= \mu_q + \rho_{qq} q_t + \sigma_{qq}\sqrt{q_t}\left(\left(1-\lambda^2\right)^{1/2}\varepsilon^q_{t+1} + \lambda\varepsilon^c_{t+1}\right) \\
\Delta c_{t+1} &= \mu_c + \rho_{cc}\Delta c_t + \rho_{cu} u_t + \sigma_{cc}\sqrt{q_t}\varepsilon^c_{t+1} \\
u_{t+1} &= \mu_u + \rho_{uu} u_t + \rho_{uc}\Delta c_t + \sigma_{uc}\sqrt{q_t}\varepsilon^c_{t+1} + \sigma_{uu}\varepsilon^u_{t+1} \\
\Delta d_{t+1} &= \Delta c_{t+1} - \delta - \Delta u_{t+1} \\
\pi_{t+1} &= \mu_\pi + \rho_\pi \pi_t + \sigma_\pi \varepsilon^\pi_{t+1}
\end{aligned} \tag{13}$$

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

$$Var_t\left[m_{t+1}\right] = \gamma^2 q_t\left[\sigma^2_{cc} + \sigma^2_{qq} - 2\sigma_{cc}\sigma_{qq}\lambda\right]$$

Consequently, increases in q_t will increase the Sharpe Ratio of all assets in the economy, and the effect will be greater the more negative is λ. If q_t and Δc_t are negatively correlated, the Sharpe ratio will increase during economic downturns (decreases in Δc_t). Note that Campbell and Cochrane essentially maximize the volatility of the pricing kernel by setting $\lambda = -1$.

3 Bond and Stock Pricing in the Moody Investor Economy

3.1 A General Pricing Model

We collect the state variables in the vector $Y_t = [q_t, \Delta c_t, u_t, \pi_t]'$. As shown in the Appendix, the dynamics of Y_t described in Equation (13) represent a simple, first-order vector autoregressive process:

$$\begin{aligned} Y_t &= \mu + AY_{t-1} + (\Sigma_F F_{t-1} + \Sigma_H)\,\varepsilon_t \\ F_t &= (\shortparallel \phi + \Phi Y_t \shortparallel)' \odot I, \end{aligned} \tag{14}$$

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, $\shortparallel \cdot \shortparallel$ denotes the non-negativity operator for a vector[6], and $\odot$ denotes the Hadamard Product.[7] 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:

$$\begin{aligned} \Sigma_F F_t \Sigma_H' &= 0 \\ \Sigma_{mF}' F_t \Sigma_{mH} &= 0 \\ \Sigma_H F_t \Sigma_{mF} &= 0 \\ \Sigma_F F_t \Sigma_{mH} &= 0 \\ \phi + \Phi Y_t &\geq 0 \end{aligned} \tag{15}$$

[6] Specifically, if v is a k-vector, then $\shortparallel v \shortparallel = \sqrt{w}$ where $w_i = \max(v_i, 0)$ for $i = 1, ..., k$.

[7] The 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 $F_t F_t' = (\phi + \Phi Y_t) \odot I$.

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 (15) 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 in the specification of q_t in Equation (4), we model f directly and bound it from below 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:

$$P_{n,t} = E_t \left[\exp \left(\hat{m}_{t+1} \right) P_{n-1,t+1} \right], \tag{16}$$

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:*

$$p_{n,t} = a_n^0 + a_n' Y_t \tag{17}$$

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

$$\begin{aligned}
a_n^0 &= a_{n-1}^0 + \left(a_{n-1} - e_\pi\right)' \mu + \mu_m + \frac{1}{2} \left(\left(\Sigma_F' \left(a_{n-1} - e_\pi\right) \right) \odot \left(\Sigma_F' \left(a_{n-1} - e_\pi\right) \right) \right)' \phi \\
&\quad + \frac{1}{2} \left(a_{n-1} - e_\pi\right)' \Sigma_H \Sigma_H' \left(a_{n-1} - e_\pi\right) + \frac{1}{2} \left(\Sigma_{mF} \odot \Sigma_{mF} \right)' \phi \\
&\quad + \frac{1}{2} \Sigma_{mH}' \Sigma_{mH} + \frac{1}{2} \sigma_m^2 + \left(a_{n-1} - e_\pi\right)' \left[\left(\Sigma_{mF}' \odot \Sigma_F \right) \phi + \Sigma_H \Sigma_{mH} \right] \\
a_n' &= \left(a_{n-1} - e_\pi\right) A + \Gamma_m' + \frac{1}{2} \left(\left(\Sigma_F' \left(a_{n-1} - e_\pi\right) \right) \odot \left(\Sigma_F' \left(a_{n-1} - e_\pi\right) \right) \right)' \Phi \\
&\quad + \frac{1}{2} \left(\Sigma_{mF} \odot \Sigma_{mF} \right)' \Phi + \left(a_{n-1} - e_\pi\right) \left[\left(\Sigma_{mF}' \odot \Sigma_F \right) \Phi \right]
\end{aligned} \tag{18}$$

where e_π is a vector such that $\pi_t = e'_\pi Y_t$ 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 models[8]. The process for the one-period short rate process, $r_t \equiv y_{1,t}$, is therefore simply $-(a^0_1 + a'_1 Y_t)$.

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

$$\begin{aligned} R^b_{n,t+1} &= \exp(a^0_{n-1} - a^0_n + a'_{n-1} Y_{t+1} - a'_n Y_t) - 1, \\ r^b_{n,t+1} &= a^0_{n-1} - a^0_n + a'_{n-1} Y_{t+1} - a'_n Y_t. \end{aligned} \tag{19}$$

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

$$V_t = E_t \left[\exp(m_{t+1}) \left(D_{t+1} + V_{t+1} \right) \right]. \tag{20}$$

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

$$PD_t = \frac{V_t}{D_t} = E_t \left\{ \sum_{n=1}^{\infty} \exp \left[\sum_{j=1}^{n} (m_{t+j} + \Delta d_{t+j}) \right] \right\}, \tag{21}$$

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

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

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

$$PD_t = \sum_{n=1}^{\infty} \exp\left(b^0_n + b'_n Y_t \right) \tag{22}$$

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

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

$$
\begin{aligned}
b_n^0 &= b_{n-1}^0 + (b_{n-1} + e_{d1})' \mu + \mu_m + \frac{1}{2} \left((\Sigma_F' (b_{n-1} + e_{d1})) \odot (\Sigma_F' (b_{n-1} + e_{d1})) \right)' \phi \\
&+ \frac{1}{2} (b_{n-1} + e_{d1})' \Sigma_H \Sigma_H' (b_{n-1} + e_{d1}) + \frac{1}{2} (\Sigma_{mF} \odot \Sigma_{mF})' \phi \\
&+ \frac{1}{2} \Sigma_{mH}' \Sigma_{mH} + \frac{1}{2} \sigma_m^2 + (b_{n-1} + e_{d1})' \left[(\Sigma_{mF}' \odot \Sigma_F) \phi + \Sigma_H \Sigma_{mH} \right] \\
b_n' &= e_{d2}' + (b_{n-1} + e_{d1})' A + \Gamma_m' + \frac{1}{2} \left((\Sigma_F' (b_{n-1} + e_{d1})) \odot (\Sigma_F' (b_{n-1} + e_{d1})) \right)' \Phi \\
&+ \frac{1}{2} (\Sigma_{mF} \odot \Sigma_{mF})' \Phi + (b_{n-1} + e_{d1}) \left[(\Sigma_{mF}' \odot \Sigma_F) \Phi \right] \qquad (23)
\end{aligned}
$$

where e_{d1} *and* e_{d2} *are selection vectors such that* $\Delta d_t = e_{d1}' Y_t + e_{d2}' Y_{t-1}$.

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:

$$
\begin{aligned}
R_{t+1}^e &= \exp(\pi_{t+1} + \Delta d_{t+1}) \left(\frac{\sum_{n=1}^{\infty} \exp\left(b_n^0 + b_n' Y_{t+1}\right) + 1}{\sum_{n=1}^{\infty} \exp\left(b_n^0 + b_n' Y_t\right)} \right) - 1 \qquad (24) \\
r_{t+1}^e &= (\pi_{t+1} + \Delta d_{t+1}) + \ln \left(\frac{\sum_{n=1}^{\infty} \exp\left(b_n^0 + b_n' Y_{t+1}\right) + 1}{\sum_{n=1}^{\infty} \exp\left(b_n^0 + b_n' Y_t\right)} \right).
\end{aligned}
$$

The only intuition immediately apparent from comparing Equations (18) and (**??**) is that the coefficient recursions look identical except for the presence of the vector $-e_\pi$ in the bond equations and e_{d1} and e_{d2} 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_{d1} and e_{d2} select dividend growth from the state variables, their 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,

$$
\begin{aligned}
r_t^{\text{real}} &= -\ln(\beta) + \gamma \left(\mu_c - \mu_q \right) + \gamma \rho_{cc} \Delta c_t + \gamma \rho_{cu} u_t \\
&+ \left[\gamma \left(1 - \rho_q \right) - \frac{1}{2} \gamma^2 \left(\sigma_{cc}^2 + \sigma_{qq}^2 - 2 \sigma_{cc} \sigma_{qq} \lambda \right) \right] q_t. \qquad (25)
\end{aligned}
$$

Hence, the real interest rate follows a three-factor model with two observed factors (consumption 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

$$r_t^{\text{real,M-P}} = -\ln(\beta) + \gamma E_t(\Delta c_{t+1}) - \frac{1}{2}\gamma^2 V_t(\Delta c_{t+1}). \tag{26}$$

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 two additional terms:

$$r_t^{\text{real}} = r_t^{\text{real,M-P}} + \gamma\left[\left(1-\rho_q\right)q_t - \mu_q\right] - \frac{1}{2}\gamma^2\left(\sigma_{qq}^2 - 2\sigma_{cc}\sigma_{qq}\lambda\right)q_t \tag{27}$$

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. Taken together, these additional terms provide sufficient channels for this economy to mitigate, in theory, 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,

$$\exp\left(-r_t^f\right) = E_t\left[\exp\left(m_{t+1} - \pi_{t+1}\right)\right]. \tag{28}$$

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.[9]:

$$r_t^f = r_t^{\text{real}} + \mu_\pi + \rho_\pi \pi_t - \frac{1}{2}\sigma_\pi^2. \tag{29}$$

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:

$$r_{t,2}^f - r_t^f = \frac{1}{2}\left(E_t\left[r_{t+1}^f\right] - r_t^f\right) + \frac{1}{2}Cov_t\left[m_{t+1}, r_{t+1}^f\right] - \frac{1}{4}Var_t\left[r_{t+1}^f\right] \tag{30}$$

The term in the middle determines the 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:

$$Cov_t\left[m_{t+1}, r_{t+1}^f\right] = v_0 + v_1 q_t \tag{31}$$

The time-variation in the term premium is entirely driven by stochastic risk aversion. Further assume that there is little movement in the conditional mean of consumption growth ($\rho_{cc} = \rho_{cu} = 0$). In this case, $v_0 = 0$ and

$$v_1 = \gamma\theta\left(\sigma_{qq}^2 - \sigma_{qq}\sigma_{cc}\lambda\right) \tag{32}$$

[9] The 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, $\exp(r_t^{\text{real}})$.

with $\theta = \gamma\left(1-\rho_q\right) - \frac{1}{2}\gamma^2\left(\sigma_{cc}^2 + \sigma_{qq}^2 - 2\sigma_c\sigma_q\lambda\right)$. Assuming λ is negative, the interpretation is straightforward. The parameter θ measures whether the precautionary savings or consumption smoothing effect dominates in the determination of interest rates. Wachter (2006) 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 (**??**), 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 consumption growth is given by

$$b_n^c = \left(b_{n-1}^c + 1\right)\rho_{cc} + \left(b_{n-1}^u - 1\right)\rho_{uc} - \gamma\rho_{cc} \tag{33}$$

Consumption growth affects equities through cash-flow and discount rate channels. In our model, dividend growth equals consumption growth minus the change in the consumption-dividend ratio. Because dividends are the cash flows of the equity shares, an increase in expected dividends should raise the price-dividend ratio. Consumption growth potentially forecasts dividend growth through two channels - future consumption growth and the future consumption-dividend ratio. This is reflected in the terms, $\left(b_{n-1}^c + 1\right)\rho_{cc}$ and $\left(b_{n-1}^u - 1\right)\rho_{uc}$. Additionally, consumption growth may forecast itself and because it is an element of the pricing kernel, this induces a discount rate effect. For example, if consumption growth is positively autocorrelated, an increase in consumption lowers expected future marginal utility. The resulting increased discount rate depresses the price-dividend ratio. This effect is represented by the term, $-\gamma\rho_{cc}$. 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 tend to be countervailing effects, it is difficult to generate much variability in price dividend ratios in such a model.

The consumption-dividend ratio effect on equity valuation is similar. The effect of the consumption dividend ratio, b_n^u is given by,

$$b_n^u = \left(b_{n-1}^c + 1\right) \rho_{cu} + \left(b_{n-1}^u - 1\right) \rho_{uu} + 1 - \gamma \rho_{cu} \tag{34}$$

The first three terms represent the effects of the consumption-dividend ratio forecasting dividends - cash flow effects - and the fourth term arises because the consumption-dividend ratio may forecast consumption growth, leading to a discount rate effect.

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:

$$\begin{aligned} b_n^q &= b_{n-1}^q \rho_{qq} \\ &\quad + \gamma \left(\rho_{qq} - 1\right) + \frac{1}{2}\gamma^2 \left(\sigma_{qq}\lambda - \sigma_{cc}\right)^2 + \frac{1}{2}\gamma^2 \sigma_{qq}^2 \left(1 - \lambda^2\right) \\ &\quad + \frac{1}{2}\left(\sigma_{cc}\left(b_{n-1}^c + 1\right) + \sigma_{uc}\left(b_{n-1}^u - 1\right) + \sigma_{qq}\lambda b_{n-1}^q\right)^2 + \frac{1}{2}\left(\sigma_{qq} b_{n-1}^q\right)^2 \left(1 - \lambda^2\right) \\ &\quad + \gamma \left(\sigma_{qq}\lambda - \sigma_{cc}\right)\left(\left(b_{n-1}^c + 1\right)\sigma_{cc} + \left(b_{n-1}^u - 1\right)\sigma_{uc} + b_{n-1}^q \sigma_{qq}\lambda\right) + \gamma\left(\sigma_{qq} b_{n-1}^q\right)^2 \left(1 - \lambda^2\right) \end{aligned} \tag{35}$$

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 term on the right hand side of Equation (35) arises only due to persistence in q_t. The second line of Equation (35) summarizes the effect of q_t on real interest rates. The first of these terms captures the intuition that if risk aversion (low surplus consumption) is high today, it is expected to be lower in the future. This induces a motive for investors to borrow against future better times, so interest rates must increase in equilibrium to discourage this borrowing, inducing a fall in the prices of long lived assets. The second and third terms on the second line of Equation (35) are precautionary savings effects. High q_t implies high uncertainty, which serves to lower rates and raise prices of long lived assets.

The third line of Equation (35) is comprised of Jensen's inequality terms, in effect reflecting an additional precautionary savings effect for assets with risky cash flows. High q_t raises the volatility of the dividend stream, and in a log-normal framework, this increases valuations.

The fourth line of Equation (35) is the most interesting because it captures the effect of the riskiness of the dividend stream on valuations, or more precisely, the effect of q_t on that riskiness.

To clean up the algebra, let us consider the direct impact of q_t (that is, excluding the b_{n-1} terms). Then the last line of Equation (35) reduces to

$$\gamma\left(\sigma_{qq}\lambda-\sigma_{cc}\right)\left(\sigma_{cc}-\sigma_{uc}\right) \tag{36}$$

Assuming that $\lambda<0$, the second term is negative. Now, if dividend growth is procyclical, covarying positively with consumption growth, then $(\sigma_{cc}-\sigma_{uc})>0$ and the overall expression in (36) is negative. Hence, in times of high risk aversion and high market volatility (high q_t), equity valuations fall.

4 Estimation and Testing Procedure

4.1 Estimation Strategy

Our economy has four state variables, which we collect in the vector Y_t. Except for q_t, 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 Z_t. That is, $Z_t=\left[\Delta c_t,u_t,\pi_t,r_t^f,spd_t,dp_t,r_t^{ex},r_t^{bx}\right]'$. Also, we let Ψ denote the structural parameters of the model:

$$\Psi=\left[\mu_c,\mu_\pi,\mu_u,\mu_q,\delta,\rho_{cc},\rho_{uc},\rho_{cu},\rho_{uu},\rho_{\pi\pi},\rho_{qq},\sigma_{cc},\sigma_{uc},\sigma_{uu},\ \sigma_{\pi\pi},\sigma_{qq},\lambda,\beta,\gamma\right]' \tag{37}$$

Throughout the estimation, we require that Ψ satisfies the conditions of Equations (15). 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 (22) and (24), 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 expression. 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 Z_t obtains,

$$Z_t = \mu^z + \Gamma^z Y_{t-1} + (\Sigma_F^z F_{t-1} + \Sigma_H^z)\, \varepsilon_t \tag{38}$$

where the coefficients superscripted with 'z' are nonlinear functions of the model parameters, Ψ. We estimate the model in a two-step GMM procedure utilizing selected conditional moments and extracting the latent state vector using the linear Kalman filter. We first describe the filtering process and then the calculation of the conditional GMM residuals and objective function. The next subsection describes the specific moments and GMM weighting matrix employed.

To filter the state vector, we represent the model in state-space form using Equation (14) as the state equation and appending Equation (38) with measurement error for the observation equation,

$$Z_t = \mu^z + \Gamma^z Y_{t-1} + (\Sigma_F^z F_{t-1} + \Sigma_H^z)\, \varepsilon_t + D v_t \tag{39}$$

where v_t is an independent standard normal measurement error innovation, and D is a diagonal matrix with the standard deviation of the measurement errors along the diagonal. It is necessary to introduce measurement error because the dimensionality of the observation equation is greater than that of the state equation (that is, the model has a stochastic singularity). To avoid estimating the measurement error variances and to keep them small, we simply fix the diagonal elements of D such that the variance of the measurement error is equal to one percent of the unconditional sample variance for each variable. Together, the state and measurement equations may be used to extract the state vector in the usual fashion using the standard linear Kalman filter (see Harvey 1989).

Given conditional (filtered) estimates for Y_t, denoted, $\widehat{Y}_t$, it is straightforward to calculate conditional moments of Z_{t+1} using Equation (39),

$$\begin{aligned} E_t\left[Z_{t+1}\right] &= \mu^z + \Gamma^z \widehat{Y}_t \\ VAR_t\left[Z_{t+1}\right] &= \left(\Sigma_F^z \widehat{F}_t + \Sigma_H^z\right)\left(\Sigma_F^z \widehat{F}_t + \Sigma_H^z\right)' + DD' \end{aligned} \tag{40}$$

where $\widehat{F}_t$ is defined analogously to Equation (14). Residuals are defined for each variable as $v_t^z = Z_t - E_{t-1}\left[Z_t\right]$.

4.2 Moment Conditions, Starting Values and Weighting Matrix

We use a total of 30 moment conditions to estimate the model parameters. They can be ordered into several groups.

$$\begin{aligned} v_t^z \times [1] \text{ for} Z_t &= \left[\Delta d_t, \Delta c_t, \pi_t, r_t^f, dp_t, spd_t, r_t^{ex}, r_t^{bx}\right] (8) \\ v_t^z \times Z_{t-1} \text{ for} Z_t &= \left[\Delta d_t, \Delta c_t, \pi_t, r_t^f, u_t\right] \ (5) \\ \left(v_t^z\right)^2 \times [1] \text{ for} Z_t &= \left[\Delta d_t, \Delta c_t, \pi_t, r_t^f, dp_t, spd_t, r_t^{ex}, r_t^{bx}, u_t\right] \ (9) \\ v_t^{z_1} \otimes v_t^{z_2} \text{ for } Z_t^1 &= \left[\Delta d_t, \Delta c_t\right], \ Z_t^2 = \left[r_t^{ex}, r_t^{bx}\right] \ (4) \\ & \left(v_t^{\Delta c} v_t^{\Delta d}\right), \left(v_t^{\Delta c} v_t^{u}\right), \left(v_t^{\Delta c} u_{t-1}\right), \left(v_t^{u} \Delta c_{t-1}\right) \ (4) \end{aligned} \tag{41}$$

The first line of (41) essentially captures the unconditional mean of the endogenous variables. 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 explicitly require the model to match the mean equity premium.

The second group uses lags of the endogenous variables as instruments to capture conditional mean dynamics for the ‘fundamental’ series and the short rate. This explicitly requires the model to address predictability (or lack thereof) of consumption and dividend growth. The third set of moments is included so that the model matches the volatility of the endogenous variables. 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. Interest rates are functions of expected marginal rates of substitution and their variability must not be excessively high to yield realistic predictions.

The fourth group captures the covariance between fundamentals and returns. These moments confront the model directly with the Cochrane-Hansen (1990) puzzle.

Finally, the fifth set is included so that the model may match the conditional dynamics between consumption and dividends. Because in Campbell and Cochrane (1999) consumption and dividends coincide, matching the conditional dynamics of these two variables is an important departure and extension.

This set of GMM residuals forms the basis of our model estimation. To optimally weight these orthogonality conditions and provide the minimization routine with good starting values, we employ a preliminary estimation which yields a consistent estimate of Ψ, denoted Ψ^1. The preliminary estimation uses only uncentered, unconditional moments of Z_t and does not require filtering of the latent state variables or a parameter dependent weighting matrix. Details of the first stage estimation are relegated to the appendix. Given Ψ^1, the residuals in (41) are calculated, and their joint spectral density at frequency zero is calculated using the Newey-West (1987) procedure. The inverse of this matrix is used as the optimal GMM weighting matrix in the main estimation stage. Note that there are 13 over-identifying restrictions and that we can use the standard J-test to assess the fit of the model.

4.3 Tests of Additional 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. In particular, we are interested in how well the model fits the bond-stock return correlation and return predictability. To test conformity of the estimated model with moments not explicitly fit in the estimation stage, we construct a GMM-based test statistic that takes into account the sampling error in estimating the parameters, Ψ. The appendix describes the exact computation.

5 Estimation Results

This section examines results from model estimation and implications for observable variables under the model.

5.1 Parameters

Table 2 reports the parameter estimates for the model. 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, neither μ_q nor μ_u are estimated, but fixed at unity and zero respectively. This is necessary for identification of the model and reduces the number of estimated parameters to 17. 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. Consumption growth shows modest serial persistence as is true in the data. The consumption-dividend ratio is quite persistent, and there is some evidence of significant feedback between (past) consumption growth and the future consumption-dividend ratio. Both inflation and stochastic risk aversion, q_t, are very persistent processes.

The volatility parameters are reported in the third column. Consumption growth is negatively correlated with the consumption-dividend ratio, but the coefficient is only significantly different from zero at about the 10 percent level. The conditional correlation between consumption growth and q_t is -0.19, and this value is statistically different from zero at about the 10 percent level. 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 17 parameters and 30 moment conditions, making the J-test a $\chi^2(13)$ under the null. The test fails to reject at the 1 percent level of significance, but rejects at the 5 percent level. The fact that all t-statistics are over 1.00 suggests that the data contain enough information to identify the parameters.

5.2 Implied Moments

Here, we assess which moments the model 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 5.2 percent on equity, which matches the data moment of 5.9 percent quite well. Standard power utility models typically do so at the cost of exorbitantly high-risk free rates, a phenomenon called the risk free rate puzzle (Weil, 1989). Our interest rate process does have a mean that is too high by 1.6 percent, but we also generate an average excess bond return of 1.1%, which is very close to the 1.0 percent data mean.

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 percent, we generate excess return volatility of 17.2 percent. What makes this especially surprising is that the model slightly undershoots the volatility of the fundamentals. That is, although they are within 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. To nevertheless generate substantial equity return volatility, 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 9.6 percent versus 8.1 percent in the data. Short rate volatility is actually within 10 basis points of the sample volatility.

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.

About 5.5 percent of the variation in excess stock returns is explained by consumption growth and the consumption dividend ratio. The bulk of the variance of returns (over 90 percent) 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[10], we are able to generate much variability in bond returns simply using a stochastic inflation process, which accounts for 80 percent of the variation. The remainder is primarily due to stochastic risk aversion, and only about 10 percent is due to consumption 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.[11]

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 (2007) and Lettau and Ludvigson (2005) for recent articles on this topic). Table 3 shows that our model matches the variance of dividend yields to within 20 basis points. Table 4 shows that the bulk of this variation comes from stochastic risk aversion and not from cash flows.

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 0.87 (versus 0.90 in the data). Additionally, the model implied term spread is a bit more persistent, at 0.83, than the data value of 0.73.

[10] An earlier unpublished version of Campbell and Cochrane (1999) relaxes this condition.

[11] In this data sample, a regression of excess bond returns on contemporaneous and lagged consumption growth and the consumption-dividend ratio yields an r-squared statistic of 0.09. A regression run on simulated data under the model reported in Table 2 yields an r-squared of 0.01.

The term structure model in this economy is affine and variation in yields is driven by four factors: consumption 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 92 percent of the total variation of the short rate, but only 83 percent 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 consumption growth explains 3 percent of the variation in short rates, it drives 13 percent of the variation in the term spread. This is natural as consumption growth is less persistent than the main driver of the short rate, inflation, making its relative weight for term spreads (which depend on expected changes in interest rates) larger (see Equation 25).

Table 3 shows that in our economy negative consumption or dividend growth shocks (recessions) are associated with lower nominal short rates and higher spreads. Such pro-cyclical interest rates and counter-cyclical spreads are consistent with conventional wisdom about interest rates, but the effects as measured relative to annual dividend and consumption growth are not very strong in the current data sample[12]. 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 positive. The model generated correlation between the short rate and dividend growth is 0.10, well within one standard deviation of the sample statistic, and the model generated correlation between the short rate and consumption growth is 0.17, just more than one standard deviation above the sample estimate. The correlation between the term spread and both consumption and dividend growth is negative as in the data, but the magnitudes are a bit too large.

Where the model has some trouble is in fitting the correlation between inflation and the term structure. Nevertheless, the variation of the term spread accounted for by inflation is almost identical to variation estimated by Ang, Bekaert and Wei (2008), whereas the contribution of inflation to the short rate variance seems slightly too large.

[12] Ang and Bekaert and Wei (2008) find that these notions better apply to real rates and find evidence that real-rates are indeed pro-cyclical.

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)). Additionally, the short rate reacts positively to consumption growth. The term spread reacts less strongly to preference shocks and also reacts negatively to an increase in consumption growth.

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 present 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 dividend and consumption growth and excess equity returns of 0.19 and 0.30 respectively, which is not significantly above the correlation in the data (which is respectively 0.09 for dividend growth and 0.19 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 both cases and closely approximately the magnitudes.

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 2.10 and its interquartile range is [1.39, 4.11]. Risk aversion is positively skewed and the 90 percentile observation equals 9.66. It has less than a 1% chance of reaching 100. Mean risk aversion is 11.07.

The bottom panel of Table 5 reports the correlation of q_t with all of our endogenous and exogenous variables. As expected, the variable is countercyclical, showing a negative correlation with both dividend and consumption growth, but somewhat weakly so. 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 small.

The top two panels of Figure 1 plot the conditionally filtered values for the latent variable, $\widehat{q}_t$, and local risk aversion, $\widehat{RA}_t = \widehat{\gamma} \exp(\widehat{q}_t)$. The model identifies the highest risk aversion following the Great Depression in the 1930's, with values briefly exceeding 50. While risk aversion generally decreased afterwards, it remained relatively high through the 1950's. Risk aversion was low during most of the 1960's and 1970's, but ramped up in the early 1980's. The stock market boom of the 1980's and 1990's was accompanied by a significant decline in risk aversion.

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. Before we do, note that Table 3 reveals that the fit of the model with respect to bond and stock market returns (last two columns and last two rows) is impressive. Of 19 moments, only two model-implied moments are outside of a two-standard error range around the sample moment.

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 0.15 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 0.28, which is less than two standard errors above the sample moment. Table 6, Panel A, presents a formal test to see how well the model fits the conditional covariance between stock and bond returns, incorporating both sampling error and parameter uncertainty (see appendix). The test fails to reject. A conditional test using the short rate, dividend yield and spread as instruments, also does not reject the model's predictions.

The bottom panel of Figure 1 plots the model implied conditional correlation between stock and bond returns over the sample period (solid blue) along with sample values of the correlation over a backward-looking rolling window of width 15 years. The low-frequency dynamics of the two series are quite similar, falling simultaneously in the late 1960s and again in the late 1990s.

The reason we overshoot the correlation on average 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, Panels B and C report on the consistency of the model with the predictability of returns in the data. 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 (2007) and Campbell and Yogo (2006).

Panel B demonstrates that 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 square brackets. 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 return regressions. 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 (2006) 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 for both stock and bond returns and also matches the fact that the coefficient is larger for the equity than for the bond return regression. In a recent paper, Buraschi and Jiltsov (2007) also find that an external habit model helps fit deviations of the Expectations Hypothesis.

In Panel C, we present alternative tests computing the model implied innovations to the return series and testing whether they are orthogonal to observable instruments. The p-values for these tests show a failure to reject in each and every case. These tests present further evidence that our model is consistent with the dynamics of expected returns.

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 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 that 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 is rejected at the 5 percent level, but still fits 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 do so. This relationship deserves further scrutiny in a model where the inflation process gets more attention.

Our article is part of a growing literature that explores the effects of stochastic risk aversion on asset price dynamics. A number of articles have stayed fairly close to the Campbell and Cochrane framework and empirically focused mainly on the term structure and deviations of the Expectations Hypothesis. These articles include Wachter (2006), Brandt and Wang (2003), who model risk aversion as a function of unexpected inflation, and Dai (2003), who constructs a model nesting internal and external habit. Other authors have explored alternative preference specifications where risk aversion varies through time. These include the regime –switching risk aversion model of Gordon and St-Amour (2000, 2004) and the preference shock model of Lettau and Wachter (2007) (who focus on explaining the value premium). Lettau and Wachter stress that it is important that there is no correlation between preference shocks in their model and fundamentals, and that an external habit model imposing a perfectly negative correlation would not work. The strength of our framework is that we remain tied to fundamentals but relax the perfect negative correlation assumption. Finally, Bekaert Engstrom and Xing (2009) try to disentangle risk aversion and economic uncertainty, which they take as the heteroskedasticity of fundamentals.

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, 2004, Luxury Goods and the Equity Premium, Journal of Finance. 59, 2959-3004.

[4] Ang, A. and G. Bekaert, 2007, Stock Return Predictability: Is it There?, Review of Financial Studies, 20, 651-707.

[5] Ang, A. and G. Bekaert, 2008, The Term Structure of Real Rates and Expected Inflation, Journal of Finance, 63, 797-849.

[6] Ameriks, J. and S. Zeldes, 2004, 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, Journal of Finance, 60, 1639-1672.

[10] Bansal, R. and M. Yaron, 2004, Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles, Journal of Finance, 59, 1481-1509.

[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., Engstrom, E., and Y. Xing, 2009, Risk, Uncertainty and Asset Prices, Journal of Financial Economics, 91, 59-82.

[13] 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.

[14] Bekaert, G., R. J. Hodrick and D. A. Marshall, 2001, Peso problem explanations for term structure anomalies, Journal of Monetary Economics 48, 241-270.

[15] Brandt, M. and K. Q. Wang, 2003, Time-Varying Risk Aversion and Unexpected Inflation, Journal of Monetary Economics, 50, 1457-1498.

[16] Brock, W., 1982, Asset Prices in a Production Economy, Scientific Essays of William Allen Brock, Elgar.

[17] Buraschi, A. and A. Jiltsov, 2007, Habit Formation and Macroeconomic Models of the Term Structure of Interest Rates, Journal of Finance, 62, 3009-3063.

[18] Campbell, J. Y., 1990, Measuring the Persistence of Expected Returns, American Economic Review 80, 43-47.

[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, 2006, Efficient Tests of Stock Return Predictability, Journal of Financial Economics, 81, 27-60.

[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] Gordon, S. and P. St. Amour, 2004, Asset Returns with State-Dependent Risk Preferences, Journal of Business and Economic Statistics, 22, 241-252.

[35] Gordon, S. and P. St. Amour, 2000, A Preference Regime Model of Bear and Bull Markets, American Economic Review, 90, 1019-1033.

[36] Hamilton, J. D., 1994, Time Series Analysis, Princeton University Press.

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

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

[39] Harvey, A., 1989, Forecasting, Structural Time Series Models, and the Kalman Filter, Cambridge University Press.

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

[41] Heaton, J., 1995, An empirical investigation of asset pricing with temporally dependent preference specifications, Econometrica 63, 681-717.

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

[43] Kumar, S., and A. Persaud, 2002, Pure Contagion and Investors' Shifting Risk Appetite, International Finance, 5, 401-436.

[44] Lettau, M. and S. Ludvigson, 2005, Expected Returns and Expected Dividend Growth, Journal of Financial Economics, 76, 583-626.

[45] Lettau, M. and J. Wachter, 2007, Why is Long-Horizon Equity Less Risky? A Duration-Based Explanation of the Value Premium, Journal of Finance, 62, 55-92.

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

[47] Lustig, H. and S.Van Nieuwerburgh, 2005, Housing Collateral, Consumption Insurance and Risk Premia: an Empirical Perspective, Journal of Finance, 60, 1167-1219.

[48] Mamaysky, H., 2002, Market Prices of Risk and Return Predictability in a Joint Stock-Bond Pricing model, Working Paper, Yale.

[49] Mankiw, G. and S. Zeldes, 1991, The Consumption of Stockholders and Nonstockholders, Journal of Financial Economics, 29, 97-112.

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

[51] Menzly, L., Santos, T., and P. Veronesi, 2004, Understanding Predictability, Journal of Political Economy, 112, 1-47.

[52] Newey, W., and K. West, 1987, A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, Econometrica, 55, 703-708.

[53] Piazzesi, M., Schneider, M., and S. Tuzel, 2007, Housing, Consumption and Asset Pricing, Journal of Financial Economics, 83, 531-569.

[54] Santos, T., and P. Veronesi, 2006, Labor Income and Predictable Stock Returns, Review of Financial Studies, 19, 1-44.

[55] 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.

[56] 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.

[57] Sun, T., 1992, Real and Nominal Interest Rates: A Discrete-Time Model and Its Continuous Time Limit, Review of Financial Studies, 5, 581-611.

[58] Wachter, J., 2006, A Consumption-Based Model of the Term Structure of Interest Rates, Journal of Financial Economics, 79, 365-399.

[59] Wei, M., 2004, Human Capital, Business Cycles, and Asset Pricing, working paper, Columbia University.

[60] 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

$$cd_t = \alpha + \delta t + \zeta \Delta cd_{t-1} + \rho cd_{t-1} + u_t \tag{42}$$

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		Preferences
$E[\Delta c]$	0.0317	ρ_{cc}	0.3713	σ_{cc}	0.0147	$\ln(\beta)$	−0.0887
	(0.0017)		(0.0926)		(0.0018)		(0.0119)
$E[\pi]$	0.0394	ρ_{cu}	−0.0085	σ_{uc}	−0.0215	γ	1.0249
	(0.0076)		(0.0052)		(0.0136)		(0.4093)
δ	−0.0052	ρ_{uc}	−2.6579	σ_{uu}	0.0602		
	(0.0046)		(0.9969)		(0.0122)		
		ρ_{uu}	0.7950	$\sigma_{\pi\pi}$	0.0172		
			(0.0653)		(0.0026)		
		$\rho_{\pi\pi}$	0.8793	σ_{qq}	0.4405		
			(0.0235)		(0.1525)		
		ρ_{qq}	0.8916	λ	−0.1892		
			(0.0290)		(0.1182)		
J-Stat(13)	24.2266						
(*pval*)	(0.0291)						

The estimated model is defined by

$$
\begin{aligned}
q_{t+1} &= \mu_q + \rho_{qq} q_t + \sigma_{qq}\sqrt{q_t}\left(\left(1-\lambda^2\right)^{1/2}\varepsilon^q_{t+1} + \lambda\varepsilon^c_{t+1}\right) \\
\Delta c_{t+1} &= \mu_c + \rho_{cc}\Delta c_t + \rho_{cu} u_t + \sigma_{cc}\sqrt{q_t}\varepsilon^c_{t+1} \\
u_{t+1} &= \mu_u + \rho_{uu} u_t + \rho_{uc}\Delta c_t + \sigma_{uc}\sqrt{q_t}\varepsilon^c_{t+1} + \sigma_{uu}\varepsilon^u_{t+1} \\
\Delta d_{t+1} &= \Delta c_{t+1} - \delta - \Delta u_{t+1} \\
\pi_{t+1} &= \mu_\pi + \rho_\pi \pi_t + \sigma_\pi \varepsilon^\pi_{t+1} \\
m_{t+1} &= \ln(\beta) - \gamma\Delta c_{t+1} + \gamma\Delta q_{t+1}
\end{aligned}
$$

The moments fit are (30 total)

$$
\begin{aligned}
v^z_t &\equiv \left(Z_t - E_{t-1}\left[Z_t\right]\right) \\
v^z_t \times [1] \text{ for} Z_t &= \left[\Delta d_t, \Delta c_t, \pi_t, r^f_t, dp_t, spd_t, r^{ex}_t, r^{bx}_t\right] (8) \\
v^z_t \times Z_{t-1} \text{ for} Z_t &= \left[\Delta d_t, \Delta c_t, \pi_t, r^f_t, u_t\right] \; (5) \\
\left(v^z_t\right)^2 \times [1] \text{ for} Z_t &= \left[\Delta d_t, \Delta c_t, \pi_t, r^f_t, dp_t, spd_t, r^{ex}_t, r^{bx}_t, u_t\right] \; (9) \\
v^{z_1}_t \otimes v^{z_2}_t \text{for } Z^1_t &= \left[\Delta d_t, \Delta c_t\right], Z^2_t = \left[r^{ex}_t, r^{bx}_t\right] \; (4) \\
& \left(v^{\Delta c}_t v^{\Delta d}_t\right), \left(v^{\Delta c}_t v^u_t\right), \left(v^{\Delta c}_t u_{t-1}\right), \left(v^u_t, \Delta c_{t-1}\right) \; (4)
\end{aligned}
$$

GMM standard errors are in parentheses. See test for a discussion of the estimation procedure. Note that the unconditional means of u_t and q_t are fixed at zero and one respectively. There are a total of 17 estimated parameters. Data are annual from 1927-2000 (74 years). See the 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	[0.037]	[0.032]	[0.039]	[0.057]*	[0.033]*	[0.004]	[0.052]	[0.011]
	0.035	0.032	0.032	0.041	0.039	0.009	0.059	0.010
	(0.016)	(0.003)	(0.006)	(0.005)	(0.002)	(0.002)	(0.024)	(0.009)
Std.	[0.084]	[0.016]	[0.036]	[0.033]	[0.012]	[0.014]	[0.172]	[0.096]
Dev.	0.124	0.022	0.042	0.032	0.014	0.013	0.197	0.081
	(0.022)	(0.005)	(0.007)	(0.004)	(0.002)	(0.001)	(0.022)	(0.007)
Auto.	[0.196]	[0.413]	[0.881]	[0.868]	[0.879]	[0.826]	[−0.057]	[−0.001]
Corr.	0.159	0.414	0.646	0.895	0.800	0.734	0.081	−0.080
	(0.164)	(0.185)	(0.160)	(0.105)	(0.286)	(0.106)	(0.162)	(0.164)

Correlations

	Δd_t	Δc_t	π_t	r_t^f	dp_t	spd_t	r_t^{ex}	r_t^{bx}
Δc_t	[0.53]							
	0.59							
	(0.14)							
π_t	[0.00]	[0.00]*						
	0.08	0.43						
	(0.21)	(0.19)						
r_t^f	[0.10]	[0.17]	[0.96]*					
	−0.01	0.01	0.42					
	(0.14)	(0.14)	(0.15)					
dp_t	[0.03]	[−0.02]*	[0.00]	[0.22]*				
	−0.07	−0.34	−0.10	−0.26				
	(0.14)	(0.15)	(0.23)	(0.11)				
spd_t	[−0.20]	[−0.37]	[−0.91]*	[−0.90]*	[0.16]			
	−0.16	−0.17	−0.33	−0.65	0.11			
	(0.14)	(0.13)	(0.14)	(0.10)	(0.11)			
r_t^{ex}	[0.19]	[0.30]	[0.04]	[0.05]	[−0.21]	[−0.18]*		
	0.09	0.19	0.01	−0.13	−0.43	0.13		
	(0.13)	(0.17)	(0.22)	(0.10)	(0.14)	(0.09)		
r_t^{bx}	[−0.06]	[−0.07]	[−0.43]	[−0.44]	[−0.04]	[0.41]*	[0.28]	
	−0.05	−0.06	−0.28	−0.34	−0.09	0.57	0.15	
	(0.08)	(0.10)	(0.11)	(0.13)	(0.10)	(0.07)	(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 100,000 periods. Dividend yield and excess equity return simulated moments are based upon the log-linear approximation described in the text. 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). Asterisks denote sample moments more than two standard errors away from the model implied value. See data appendix for additional data construction notes.

Table 4: Variance Decomposition Under the Moody Investor Model

	Δc_t	u_t	π_t	q_t	Δc_{t-1}	u_{t-1}	π_{t-1}	q_{t-1}
r_t^f	0.0321	0.0040	0.9188	0.0451	0.0000	0.0000	0.0000	0.0000
dp_t	0.0003	0.0902	0.0000	0.9095	0.0000	0.0000	0.0000	0.0000
spd_t	0.1273	0.0050	0.8300	0.0377	0.0000	0.0000	0.0000	0.0000
r_t^{ex}	0.0502	0.0049	0.0099	0.4264	0.0042	0.0069	0.0000	0.4976
r_t^{bx}	0.0100	0.0902	0.8083	0.0280	-0.0017	-0.0023	0.0000	0.1527

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

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

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: Properties of q_t and Risk Aversion Under the Moody Investor Model

RA_t Percentile	1%	5%	10%	25%	50%	75%	90%	95%	99%	mean
	0.90	1.05	1.13	1.39	2.10	4.11	9.66	18.4	82.0	11.07

q_t Correlation	Δd_t	Δc_t	π_t	r_t^f	dp_t	spd_t	r_t^{ex}	r_t^{bx}
	-0.06	-0.15	0.00	0.21	0.93	0.20	-0.21	-0.04

This table presents simulated risk aversion moments under the Moody Investor Model and correlations with q_t and observable variables. The system was simulated for 100,000 periods using the data generating process and point estimates from Table 2. Risk aversion is calculated as

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

Table 6: Tests of Additional Moments

Panel A: Stock-Bond Covariance Orthogonality Tests

Moment(s)	$p-val$
$u_t^{rex}u_t^{bex} - C_{t-1}\left[r_t^{ex}, r_t^{bx}\right]$	(0.34)
$\left(u_t^{rex}u_t^{bex} - C_{t-1}\left[r_t^{ex}, r_t^{bx}\right]\right) \otimes \overrightarrow{z}_{t-1}$	(0.77)

Panel B: Return Conditional Mean Orthogonality Tests

Moment(s)	$p-val$	Moment(s)	$p-val$	joint $p-val$
$u_t^{rex}r_{t-1}^f$	(0.36)	$u_t^{rbx}r_{t-1}^f$	(0.71)	(0.65)
$u_t^{rex}dp_{t-1}$	(0.56)	$u_t^{rbx}dp_{t-1}$	(0.25)	(0.47)
$u_t^{rex}spd_{t-1}$	(0.59)	$u_t^{rbx}spd_{t-1}$	(0.36)	(0.65)
$u_t^{rex} \otimes \overrightarrow{z}_{t-1}$	(0.67)	$u_t^{rbx} \otimes \overrightarrow{z}_{t-1}$	(0.48)	(0.77)

Panel C: Return Predictability Univariate Slope Coefficients

	r_{t-1}^f	dp_{t-1}	spd_{t-1}	$\left(dp_{t-1} - r_{t-1}^f\right)$
r_t^{ex}	[0.2365]	[3.2598]	[0.7057]	[0.1889]
	-0.7999	1.6748	3.5705	0.8184
	(0.5992)	(1.4016)	(1.4603)	(0.4587)
r_{t-1}^{bx}	[0.0927]	[1.1975]	[0.2332]	[0.0634]
	-0.0781	0.2121	2.1029	0.0834
	(0.2496)	(0.5828)	(0.5846)	(0.2107)

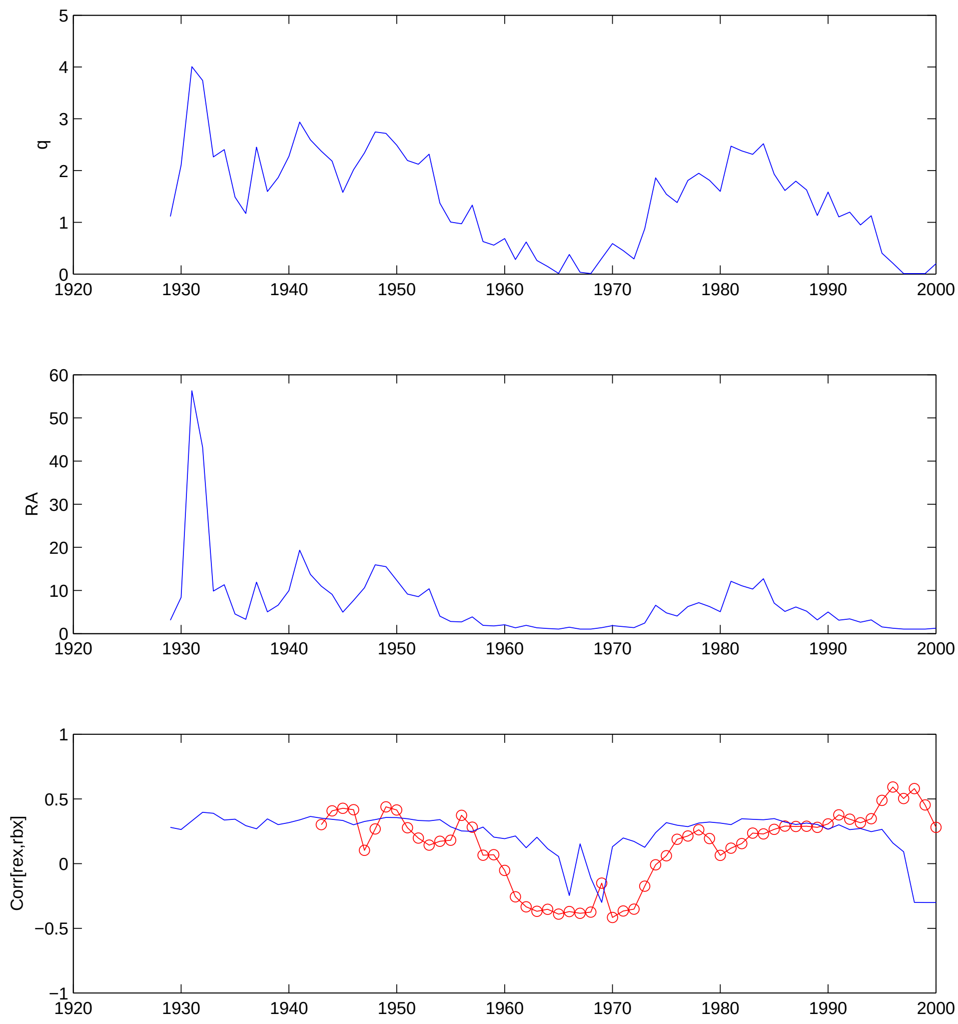
This table reports the covariance and predictability performance of the Moody Investor model. In both panels, u_t^x denotes $(x_t - E_{t-1}[x_t])$ where x_t is an observable variable. The conditional expectation is that implied by the model. $\overrightarrow{z}_{t-1} = \left[r_{t-1}^f, dp_{t-1}, spd_{t-1}\right]$

In Panel A, we test moments corresponding to the unconditional and conditional covariance between stock and bond returns. $C_{t-1}\left[r_t^{ex}, r_t^{bx}\right]$ denotes the conditional covariance implied by the model.

In panel B, we test moments capturing the conditional mean of stock and bond excess returns: the conditional risk premiums. The columns labeled 'p-val' report GMM based orthogonality tests of the corresponding moment(s) condition. The final column reports a joint test for the moments across the rows. Data are annual from 1927-2000, (74 years). See the data appendix for additional data construction notes.

In Panel C, we present slope coefficients from simulated univariate return predictability regressions using various instruments under the model (top number in cell with square brackets) and the sample coefficients in the data (second and third numbers are data slope coefficients with standard errors in parentheses)

Figure 1: Filtered Conditional Moments



This figure plots various filtered series under the Moody Investor model and the point estimates from Table2. Each of the plotted series is a function of q_t alone, conditional on the model parameters.

The first two frames plot the filtered values for the latent state variable q_t and risk aversion, $RA_t = \gamma \exp(q_t)$. The remaining frame plots the model implied conditional correlation of excess stock and bonds returns (blue) and 15-year-rolling-window realized correlations between stock and bond returns (red, circles).

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:

$$\begin{aligned} RET_t^m &= \sum_{n=1}^{N} ret_{i,t} \cdot \frac{(prc_{i,t-1} \cdot shrout_{i,t-1})}{MCAP_{t-1}^m} \\ MCAP_t^m &= \sum_{n=1}^{N} prc_{i,t} \cdot shrout_{i,t} \end{aligned} \tag{43}$$

   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:

$$r_t^{ex} \equiv \ln(1 + RET_t) - r_{t-1}^f \tag{44}$$

   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,

$$r_t^{bx} \equiv \ln(1 + LTBR_t) - r_{t-1}^f \tag{45}$$

   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:

$$\begin{aligned}\Delta d_t &\equiv \ln(DIV_t) - \ln(DIV_{t-1}) - \pi_t \\ DIV_t &= \sum_{month=1}^{12}\left(\sum_{n=1}^{N}(ret_{i,t} - retx_{i,t})(prc_{i,t-1} \cdot shrout_{i,t-1})\right) \qquad (46)\end{aligned}$$

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:

$$spd_t \equiv \ln(1 + LTBY_t) - r_t^f \qquad (47)$$

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:

$$\pi_t \equiv \ln(1 + INFL_t) \qquad (48)$$

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

$$C_t^{WT} = PART_t. \cdot \left(C_t^{AG} - C_t^{FT}\right) + C_t^{LX} \qquad (49)$$

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

(50)

From these data, an interpolated participation rate, $PART_t$ was calculated by the authors as $\left(1+\exp\left[-\left(-62.0531+0.03118\cdot YEAR_t\right)\right]\right)^{-1}$, 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:

$$\Delta c_t = \ln\left(C_t\right) - \ln\left(C_{t-1}\right) - \pi_t \tag{51}$$

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:

$$\begin{aligned} dp_t &\equiv \ln\left(1+\frac{DIV_t}{MCAP_t}\right) \\ MCAP_t &= MCAP_{t,DEC}^{m} \end{aligned} \tag{52}$$

where DIV_t is defined above and $MCAP^m_{t,DEC}$ corresponds to the December value of $MCAP^m_t$ for each year.

8. r^f_t. The short term risk free rate is defined as:

$$r^f_t \equiv \ln(1 + STBY_t) \tag{53}$$

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^f_t 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,

$$\begin{aligned} Y_t &= \mu + AY_{t-1} + (\Sigma_F F_{t-1} + \Sigma_H)\, \varepsilon_t \\ F_t &= (\| \phi + \Phi Y_t \|)' \odot I, \end{aligned} \tag{54}$$

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,

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

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

$$\begin{aligned}
\Sigma_F F_t \Sigma_H' &= 0 \\
\Sigma_{mF}' F_t \Sigma_{mH} &= 0 \\
\Sigma_H F_t \Sigma_{mF} &= 0 \\
\Sigma_F F_t \Sigma_{mH} &= 0 \\
\phi + \Phi Y_t &\geq 0
\end{aligned} \tag{55}$$

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:

$$E_t\left[\exp\left(a + c_1' Y_{t+1} + d' Y_t + m_{t+1}^r\right)\right] = \exp\left(g_0 + g' Y_t\right) \tag{56}$$

Proof. By lognormality,

$$E_t\left[\exp\left(a + c' Y_{t+1} + m_{t+1}^r\right)\right] = \exp\left(\begin{array}{c} a + c' E_t\left[Y_{t+1}\right] + E_t\left[m_{t+1}^r\right] \\ +\frac{1}{2}\left(V_t\left[c' Y_{t+1}\right] + V_t\left[m_{t+1}^r\right] + 2C_t\left[c' Y_{t+1}, m_{t+1}^r\right]\right) \end{array}\right) \tag{57}$$

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

1.
$$c' E_t\left[Y_{t+1}\right] = c'\left(\mu + A Y_t\right) \tag{58}$$

2.
$$E_t\left[m_{t+1}^r\right] = \mu_m + \Gamma_m' Y_t \tag{59}$$

3.

$$
\begin{aligned}
VAR_t\left[c'Y_{t+1}\right] &= c'\left(\Sigma_F F_t + \Sigma_H\right)\left(\Sigma_F F_t + \Sigma_H\right)' c \\
&= c'\Sigma_F F_t F_t' \Sigma_F' c + c'\Sigma_H \Sigma_H' c \\
&= \left(\left(\Sigma_F' c\right) \odot \left(\Sigma_F' c\right)\right)' \left(\phi + \Phi Y_t\right) + c'\Sigma_H \Sigma_H' c \qquad (60)
\end{aligned}
$$

where the second line uses restrictions in Equation (55) and the third line follows from properties of the $\odot$ operator.

4.

$$
\begin{aligned}
VAR_t\left[m_{t+1}^r\right] &= \left(\Sigma_{mF}' F_t + \Sigma_{mH}'\right)\left(\Sigma_{mF}' F_t + \Sigma_{mH}'\right)' \\
&= \Sigma_{mF}' F_t F_t' \Sigma_{mF} + \Sigma_{mH}' \Sigma_{mH} \\
&= \left(\Sigma_{mF} \odot \Sigma_{mF}\right)' \left(\phi + \Phi Y_t\right) + \Sigma_{mH}' \Sigma_{mH} \qquad (61)
\end{aligned}
$$

where the second line uses restrictions in Equation (55) and the third line follows from properties of the $\odot$ operator.

5.

$$
\begin{aligned}
COV_t\left[c'Y_{t+1}, m_{t+1}^r\right] &= c'\left[\left(\Sigma_F F_{t-1} + \Sigma_H\right)\left(\Sigma_{mF}' F_t + \Sigma_{mH}'\right)'\right] \\
&= c'\left[\Sigma_F F_t F_t' \Sigma_{mF} + \Sigma_H \Sigma_{mH}\right] \\
&= c'\left[\left(\Sigma_{mF}' \odot \Sigma_F\right)\left(\phi + \Phi Y_t\right) + \Sigma_H \Sigma_{mH}'\right] \qquad (62)
\end{aligned}
$$

where the second line uses restrictions in Equation (55) and the third line follows from properties of the $\odot$ operator.

Substituting,

$$
\begin{aligned}
& E_t\left[\exp\left(a + c'Y_{t+1} + m^r_{t+1}\right)\right] \\
= \quad & \exp\left(\begin{array}{c}
a + c'\left(\mu + AY_t\right) + \mu_m + \Gamma'_m Y_t \\
+\frac{1}{2}\left(\left(\Sigma'_F c\right)\odot\left(\Sigma'_F c\right)\right)'\left(\phi + \Phi Y_t\right) + \frac{1}{2}c'\Sigma_H\Sigma'_H c \\
+\frac{1}{2}\left(\Sigma_{mF}\odot\Sigma_{mF}\right)'\left(\phi + \Phi Y_t\right) + \frac{1}{2}\Sigma_{mH}\Sigma'_{mH} + \frac{1}{2}\sigma^2_m \\
+c'\left[\left(\Sigma'_{mF}\odot\Sigma_F\right)\left(\phi + \Phi Y_t\right) + \Sigma_H\Sigma'_{mH}\right]
\end{array}\right)
\end{aligned}
\tag{63}
$$

Evidently,

$$
\begin{aligned}
g_0 \quad = \quad & a + c'\mu + \mu_m + \frac{1}{2}\left(\left(\Sigma'_F c\right)\odot\left(\Sigma'_F c\right)\right)'\phi \\
& + \frac{1}{2}c'\Sigma_H\Sigma'_H c + \frac{1}{2}\left(\Sigma_{mF}\odot\Sigma_{mF}\right)'\phi + \frac{1}{2}\Sigma'_{mH}\Sigma_{mH} + \frac{1}{2}\sigma^2_m + c'\left[\left(\Sigma'_{mF}\odot\Sigma_F\right)\phi + \Sigma_H\Sigma_{mH}\right] \\
g' \quad = \quad & d' + c'A + \Gamma'_m + \frac{1}{2}\left(\left(\Sigma'_F c\right)\odot\left(\Sigma'_F c\right)\right)'\Phi + \frac{1}{2}\left(\Sigma_{mF}\odot\Sigma_{mF}\right)'\Phi + c'\left[\left(\Sigma'_{mF}\odot\Sigma_F\right)\Phi\right]
\end{aligned}
\tag{64}
$$

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,

$$
\exp\left(r^f_t\right) = \left(E_t\left[\exp\left(m^r_{t+1} - e'_\pi Y_{t+1}\right)\right]\right)^{-1} \tag{65}
$$

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

$$
r^f_t = -a^0_1 - a'_1 Y_t \tag{66}
$$

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,

$$\begin{aligned} r_t^f &= -a_1^0 - a_1' Y_t \\ p_{1,t} &= a_1^0 + a_1' Y_t \end{aligned} \tag{67}$$

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,

$$P_{n,t} = E_t \left[M_{t+1}^N P_{n-1,t+1} \right]. \tag{68}$$

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

$$P_{n-1,t} = \exp \left(a_{n-1}^0 + a_{n-1}' Y_t \right). \tag{69}$$

Then, leading this expression by one period and substituting it into the recursive relation, (68), we have,

$$\begin{aligned} P_{n,t} &= E_t \left[\exp \left(m_{t+1}^N + p_{n-1,t+1} \right) \right] \\ &= E_t \left[\exp \left(m_{t+1}^r - e_\pi' Y_{t+1} + a_{n-1}^0 + a_{n-1}' Y_{t+1} \right) \right] \\ &= \exp(a_n^0 + a_n' Y_t) \end{aligned} \tag{70}$$

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

$$\begin{aligned} a_n^0 &= a_{n-1}^0 + (a_{n-1} - e_\pi)' \mu + \mu_m + \frac{1}{2} \left(\left(\Sigma_F' (a_{n-1} - e_\pi) \right) \odot \left(\Sigma_F' (a_{n-1} - e_\pi) \right) \right)' \phi \\ &\quad + \frac{1}{2} (a_{n-1} - e_\pi)' \Sigma_H \Sigma_H' (a_{n-1} - e_\pi) + \frac{1}{2} (\Sigma_{mF} \odot \Sigma_{mF})' \phi \\ &\quad + \frac{1}{2} \Sigma_{mH}' \Sigma_{mH} + \frac{1}{2} \sigma_m^2 + (a_{n-1} - e_\pi)' \left[(\Sigma_{mF}' \odot \Sigma_F) \phi + \Sigma_H \Sigma_{mH} \right] \\ a_n' &= (a_{n-1} - e_\pi) A + \Gamma_m' + \frac{1}{2} \left(\left(\Sigma_F' (a_{n-1} - e_\pi) \right) \odot \left(\Sigma_F' (a_{n-1} - e_\pi) \right) \right)' \Phi \\ &\quad + \frac{1}{2} (\Sigma_{mF} \odot \Sigma_{mF})' \Phi + (a_{n-1} - e_\pi) \left[(\Sigma_{mF}' \odot \Sigma_F) \Phi \right] \end{aligned} \tag{71}$$

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_{d1} and e_{e2} be the two selection vectors such that $\Delta d_t = e'_{d1} Y_t + e'_{d2} Y_{t-1}$. The price dividend ratio is given by

$$\begin{aligned} \frac{P_t}{D_t} &= E_t \sum_{n=1}^{\infty} \exp\left(\sum_{j=1}^{n} m^r_{t+j} + \Delta d_{t+j} \right) \\ &= E_t \sum_{n=1}^{\infty} \exp\left(\sum_{j=1}^{n} m^r_{t+j} + e'_{d1} Y_{t+j} + e'_{d2} Y_{t+j-1} \right) \\ &= \sum_{n=1}^{\infty} q^0_{n,t} \end{aligned} \tag{72}$$

where $q^0_{n,t} \equiv E_t \exp\left(\sum_{j=1}^{n} m^r_{t+j} + e'_{d1} Y_{t+j} + e'_{d2} Y_{t+j-1} \right)$ are scalars. We will prove that $q^0_{n,t} = \exp\left(b^0_n + b'_n Y_t \right)$ where b^0_n (scalar) and b'_n (k-vectors) are defined below. The proof is accomplished by induction. Consider $q^0_{1,t}$:

$$q^0_{1,t} = E_t \left(\exp\left(m^r_{t+j} + e'_{d1} Y_{t+j} + e'_{d2} Y_{t+j-1} \right) \right) \tag{73}$$

By Lemma 1,

$$q^0_{1,t} = \exp\left(b^0_1 + b'_1 Y_t \right) \tag{74}$$

where b^0_1 and b'_1 are given by Lemma 1. Next, suppose that $q^0_{n-1,t} = \exp\left(b^0_{n-1} + b'_{n-1} Y_t \right)$. Then rearrange $q^0_{n,t}$ as follows.

$$\begin{aligned}
q^0_{n,t} &= E_t \exp\left(\sum_{j=1}^{n} m^r_{t+j} + e'_{d1}Y_{t+j} + e'_{d2}Y_{t+j-1}\right) \\
&= E_t E_{t+1}\left\{\exp\left(m^r_{t+1} + e'_{d1}Y_{t+1} + e'_{d2}Y_t\right)\exp\left(\sum_{j=1}^{n-1} m^r_{t+j+1} + e'_{d1}Y_{t+j+1} + e'_{d2}Y_{t+j}\right)\right\} \\
&= E_t\left\{\exp\left(m^r_{t+1} + e'_{d1}Y_{t+1} + e'_{d2}Y_t\right)E_{t+1}\exp\left(\sum_{j=1}^{n-1} m^r_{t+j+1} + e'_{d1}Y_{t+j+1} + e'_{d2}Y_{t+j}\right)\right\} \\
&= E_t\left\{\exp\left(m^r_{t+1} + e'_{d1}Y_{t+1} + e'_{d2}Y_t\right)q^0_{n-1,t+1}\right\} \\
&= E_t\left\{\exp\left(m^r_{t+1} + e'_{d1}Y_{t+1} + e'_{d2}Y_t\right)\exp\left(b^0_{n-1} + b'_{n-1}Y_{t+1}\right)\right\} \\
&= E_t\left\{\exp\left(b^0_{n-1} + m^r_{t+1} + \left(e_d + b_{n-1}\right)'Y_{t+1} + e'_{d2}Y_t\right)\right\} \\
&= \exp\left(b^0_n + b'_nY_t\right) \qquad (75)
\end{aligned}$$

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

$$\begin{aligned}
b^0_n &= b^0_{n-1} + \left(b_{n-1} + e_{d1}\right)'\mu + \mu_m + \frac{1}{2}\left(\left(\Sigma'_F\left(b_{n-1} + e_{d1}\right)\right)\odot\left(\Sigma'_F\left(b_{n-1} + e_{d1}\right)\right)\right)'\phi \\
&\quad + \frac{1}{2}\left(b_{n-1} + e_{d1}\right)'\Sigma_H\Sigma'_H\left(b_{n-1} + e_{d1}\right) + \frac{1}{2}\left(\Sigma_{mF}\odot\Sigma_{mF}\right)'\phi \\
&\quad + \frac{1}{2}\Sigma'_{mH}\Sigma_{mH} + \frac{1}{2}\sigma^2_m + \left(b_{n-1} + e_{d1}\right)'\left[\left(\Sigma'_{mF}\odot\Sigma_F\right)\phi + \Sigma_H\Sigma_{mH}\right] \\
b'_n &= e'_{d2} + \left(b_{n-1} + e_{d1}\right)'A + \Gamma'_m + \frac{1}{2}\left(\left(\Sigma'_F\left(b_{n-1} + e_{d1}\right)\right)\odot\left(\Sigma'_F\left(b_{n-1} + e_{d1}\right)\right)\right)'\Phi \\
&\quad + \frac{1}{2}\left(\Sigma_{mF}\odot\Sigma_{mF}\right)'\Phi + \left(b_{n-1} + e_{d1}\right)\left[\left(\Sigma'_{mF}\odot\Sigma_F\right)\Phi\right] \qquad (76)
\end{aligned}$$

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

$$\begin{aligned}
q_{t+1} &= \mu_q + \rho_{qq} q_t + \sigma_{qq}\sqrt{q_t}\left(\left(1-\lambda^2\right)^{1/2}\varepsilon^q_{t+1} + \lambda\varepsilon^c_{t+1}\right) \\
\Delta c_{t+1} &= \mu_c + \rho_{cc}\Delta c_t + \rho_{cu} u_t + \sigma_{cc}\sqrt{q_t}\varepsilon^c_{t+1} \\
u_{t+1} &= \mu_u + \rho_{uu} u_t + \rho_{uc}\Delta c_t + \sigma_{uc}\sqrt{q_t}\varepsilon^c_{t+1} + \sigma_{uu}\varepsilon^u_{t+1} \\
\Delta d_{t+1} &= \Delta c_{t+1} - \delta - \Delta u_{t+1} \\
\pi_{t+1} &= \mu_\pi + \rho_\pi \pi_t + \sigma_\pi \varepsilon^\pi_{t+1} \qquad (77)
\end{aligned}$$

$$m_{t+1} = \ln(\beta) - \gamma\Delta c_{t+1} + \gamma\left(q_{t+1} - q_t\right) \qquad (78)$$

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

$$\begin{aligned}
\mu &= \begin{bmatrix} \mu_q \\ \mu_c \\ \mu_u \\ \mu_\pi \end{bmatrix}, A = \begin{bmatrix} \rho_{qq} & 0 & 0 & 0 \\ 0 & \rho_{cc} & \rho_{cu} & 0 \\ 0 & \rho_{uc} & \rho_{uu} & 0 \\ 0 & 0 & 0 & \rho_\pi \end{bmatrix}, \Sigma_H = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_u & 0 \\ 0 & 0 & 0 & \sigma_\pi \end{bmatrix} \\
\Sigma_F &= \begin{bmatrix} \left(1-\lambda^2\right)^{1/2}\sigma_q & \lambda\sigma_q & 0 & 0 \\ 0 & \sigma_{cc} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \phi = . \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \Phi = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \qquad (79)
\end{aligned}$$

where $\sigma_{qq} \equiv \frac{1}{f_q}\sqrt{\mu_q\left(1-\rho_{qq}^2\right)}$. Note that in this formulation, f_q is the ratio of the unconditional mean to the unconditional variance of q_t. It is estimated directly.

The pricing kernel for the model can be written as:

$$\begin{aligned}
m_{t+1} &= \ln(\beta) - \gamma\Delta c_{t+1} + \gamma\left(q_{t+1} - q_t\right) \\
&= \ln(\beta) - \gamma\begin{pmatrix} -\left(\mu_c + \rho_{cc}\Delta c_t + \rho_{cu} u_t + \sigma_{cc}\sqrt{q_t}\varepsilon^c_{t+1}\right) + ... \\ \mu_q + \left(\rho_{qq} - 1\right) q_t + \sigma_{qq}\sqrt{q_t}\left(\left(1-\lambda^2\right)^{1/2}\varepsilon^q_{t+1} + \lambda\varepsilon^c_{t+1}\right) \end{pmatrix} \qquad (80)
\end{aligned}$$

We can now read off the pricing kernel matrices:

$$\mu_m = \left(\ln(\beta) - \gamma\mu_c + \gamma\mu_q\right) \qquad \Gamma_m = \gamma \begin{bmatrix} \rho_{qq} - 1 \\ -\rho_{cc} \\ -\rho_{cu} \\ 0 \end{bmatrix} \qquad \Sigma_{mH} = 0$$

$$\Sigma_{mF} = \gamma \begin{bmatrix} \left(1-\lambda^2\right)^{1/2} \sigma_{qq} \\ -\sigma_{cc} + \lambda\sigma_{qq} \\ 0 \\ 0 \end{bmatrix}. \tag{81}$$

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

$$\begin{aligned}
\Delta d_{t+1} &= \mu_d + \rho_d \Delta d_t + \rho_{dq} q_t + \sigma_d \varepsilon_{t+1}^d + \kappa_1 \sqrt{q_t} \varepsilon_{t+1}^q \\
\Delta c_{t+1} &= \mu_c + \rho_c \Delta c_t + \rho_{cq} q_t + \sigma_{cd} \varepsilon_{t+1}^d + \sigma_c \varepsilon_{t+1}^c + \kappa_2 \sqrt{q_t} \varepsilon_{t+1}^q \\
\pi_{t+1} &= \mu_\pi + \rho_\pi \pi_t + \sigma_\pi \varepsilon_{t+1}^\pi \\
q_{t+1} &= \mu_q + \rho_q q_t + \frac{1}{f} \sqrt{\mu_q \left(1-\rho_q^2\right)} \sqrt{q_t} \varepsilon_{t+1}^q \\
m_{t+1}^r &= \ln(\beta) - \gamma \Delta c_{t+1} + \gamma \Delta q_{t+1}
\end{aligned} \tag{82}$$

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 (72), we see that the price dividend ratio is given by

$$\begin{aligned} \frac{P_t}{D_t} &= \sum_{n=1}^{\infty} q_{n,t}^0 \\ &= \sum_{n=1}^{\infty} \exp\left(b_n^0 + b_n' Y_t\right) \end{aligned} \tag{83}$$

and the coefficient sequences, $\left\{b_n^0\right\}_{n=1}^{\infty}$ 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 $\overline{Y}$, the unconditional mean of Y_t. Let

$$\overline{q}_n^0 = \exp\left(b_n^0 + b_n' \overline{Y}\right) \tag{84}$$

and note that

$$\frac{\partial}{\partial Y_t}\left(\sum_{n=1}^{\infty} q_{n,t}^0\right) = \sum_{n=1}^{\infty} \frac{\partial}{\partial Y_t} q_{n,t}^0 = \sum_{n=1}^{\infty} q_{n,t}^0 \cdot b_n' \tag{85}$$

Approximating,

$$\begin{aligned} pd_t &\simeq \ln\left(\sum_{n=1}^{\infty} \overline{q}_n^0\right) + \frac{1}{\sum_{n=1}^{\infty} \overline{q}_n^0}\left(\sum_{n=1}^{\infty} \overline{q}_n^0 \cdot b_n'\right)\left(Y_t - \overline{Y}\right) \\ &= d_0 + d' Y_t \end{aligned} \tag{86}$$

where d_0 and d' are implicitly defined. Similarly,

$$\begin{aligned} gpd_t &\equiv \ln\left(1 + \frac{P_t}{D_t}\right) \simeq \ln\left(1 + \sum_{n=1}^{\infty} \overline{q}_n^0\right) + \frac{1}{1 + \sum_{n=1}^{\infty} \overline{q}_n^0}\left(\sum_{n=1}^{\infty} \overline{q}_n^0 \cdot b_n'\right)\left(Y_t - \overline{Y}\right) \\ &= h_0 + h' Y_t \end{aligned} \tag{87}$$

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

$$dp_t \equiv \ln\left(1 + \frac{D_t}{P_t}\right) = gpd_t - pd_t \tag{88}$$

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,

$$\begin{aligned} r_{t+1}^{x} &= -r_t^f - pd_t + gd_{t+1} + \pi_{t+1} + gpd_{t+1} \\ &\sim (h_0 - d_0) + (e_d' + e_\pi' + h')\, Y_{t+1} + \left(-e_{rf}' + -d'\right) Y_t \\ &= r_0 + r_1' Y_{t+1} + r_2' Y_t \end{aligned} \tag{89}$$

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	0.0363	0.0368	0.0543	0.0548
std. dev.	0.0121	0.0124	0.1704	0.1603
correlation	0.9948		0.9665	

C.4 Analytic Moments of Y_t and Z_t

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

$$\begin{aligned} Y_t &= \mu + AY_{t-1} + (\Sigma_F F_{t-1} + \Sigma_H)\, \varepsilon_t \\ F_t &= sqrt(diag(\phi + \Phi Y_t)) \end{aligned} \tag{90}$$

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

$$\begin{aligned}
\overline{Y_t} &= \left(I_k - A\right)^{-1}\mu \\
vec\left(\overline{Y_t Y_t'}\right) &= \left(I_{k^2} - A \otimes A\right)^{-1} \cdot vec\left(\mu\mu' + \mu\overline{Y_t}'A' + A\overline{Y_t}\mu' + \Sigma_F \overline{F_t^2}\Sigma_F' + \Sigma_H \Sigma_H'\right) \\
vec\left(\overline{Y_t Y_{t-1}'}\right) &= \left(I_{k^2} - A \otimes A\right)^{-1} \cdot vec\left(\mu\mu' + \mu\overline{Y_t}'A' + A\overline{Y_t}\mu' + A\left(\Sigma_F \overline{F_t^2}\Sigma_F' + \Sigma_H \Sigma_H'\right)\right)
\end{aligned} \tag{91}$$

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

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

$$Z_t = \mu^w + \Gamma^w Y_{t-1} + \left(\Sigma_F^w F_{t-1} + \Sigma_H^w\right)\varepsilon_t \tag{92}$$

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

$$\begin{aligned}
\overline{Z_t} &= \mu^w + \Gamma^w \overline{Y_t} \\
\overline{Z_t Z_t} &= \mu^w \mu^{w\prime} + \mu^w \overline{Y_t}'\Gamma^{w\prime} + \Gamma^w \overline{Y_t}\mu^{w\prime} + \Gamma^w \overline{Y_t Y_t'}\Gamma^{w\prime} + \Sigma_F^w \overline{F_t^2}\Sigma_F^{w\prime} + \Sigma_H^w \Sigma_H^{w\prime} \\
\overline{Z_t Z_t} &= \mu^w \mu^{w\prime} + \mu^w \overline{Y_t}'\Gamma^{w\prime} + \Gamma^w \overline{Y_t}\mu^{w\prime} + \Gamma^w \overline{Y_t Y_{t-1}'}\Gamma^{w\prime} + \Gamma^w\left(\Sigma_F \overline{F_t^2}\Sigma_F^{w\prime} + \Sigma_H \Sigma_H^{w\prime}\right)
\end{aligned} \tag{93}$$

It remains to demonstrate that the observable series used in estimation obey Equation (92). This is trivially true for elements of Z_t which are also elements of Y_t such as Δd_t, Δc_t, π_t. Using Equations (66), (89) and (86), it is apparent that r_t^f, dp_t and r_t^x satisfy Equation (92) as well.

C.5 Test of Additional Moments

To test conformity of the estimated model with moments not explicitly fit in the estimation stage, the a GMM based statistic is constructed that takes into account the sampling error in estimating the parameters, Ψ. Let $g_{2T}\left(\Psi_0, X_t\right)$ be the sample mean of the restrictions we wish to test. By the Mean Value Theorem,

$$g_{2T}\left(\widehat{\Psi}\right) \overset{a.s.}{=} g_{2T}\left(\Psi_0\right) + D_{2T} \cdot \left(\widehat{\Psi} - \Psi_0\right) \tag{94}$$

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

$$\left(\widehat{\Psi} - \Psi_0\right) \overset{a.s.}{=} -\left(A_{11} D_{1T}\right)^{-1} A_{11} g_{1T}\left(\Psi_0\right) \tag{95}$$

with

$$D_{1T} = \frac{\partial g_{1T}\left(\Psi_0\right)}{\partial \Psi} \tag{96}$$

$$A_{11} = D_{1T}' S_{11}^{-1} \tag{97}$$

Substituting,

$$g_{2T}\left(\widehat{\Psi}\right) \overset{a.s.}{=} L g_T\left(\Psi_0\right) \tag{98}$$

where

$$L = \left[-D_{2T} \cdot \left(A_{11} D_{1T}\right)^{-1} A_{11}',\ I\right] \tag{99}$$

$$g_T\left(\Psi_0\right) = \left[g_{1T}\left(\Psi_0\right)',\ g_{2T}\left(Z_t; \Psi_0\right)'\right]' \tag{100}$$

Since $\sqrt{T} g_T\left(\Psi_0\right) \rightarrow N(0, S)$ 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,

$$T g_{2T}\left(\widehat{\Psi}\right)\left[L S L'\right]^{-1} g_{2T}\left(\widehat{\Psi}\right) \tag{101}$$

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

D Unconditional Parameter Estimation (First Stage)

Collect all the measurable variables of interest, the three observable state variables and the five endogenous variables in the vector Z_t. Also, we let Ψ denote the structural parameters of the model:

$$\Psi = \left[\mu_c, \mu_\pi, \mu_u, \mu_q, \delta, \rho_{cc}, \rho_{uc}, \rho_{cu}, \rho_{uu}, \rho_{\pi\pi}, \rho_{qq}, \sigma_{cc}, \sigma_{uc}, \sigma_{uu}, \sigma_{\pi\pi}, \sigma_{qq}, \lambda, \beta, \gamma\right]' \tag{102}$$

Applying the log-linear approximation of Appendix C.2, the following property of Z_t obtains,

$$Z_t = \mu^z + \Gamma^z Y_{t-1} + \left(\Sigma_F^z F_{t-1} + \Sigma_H^z\right)\varepsilon_t \tag{103}$$

where the coefficients superscripted with 'z' 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 $X(Z_t)$ be a vector valued function of Z_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 Z_t. Using Equation (38), we can also derive the analytic solutions for uncentered moments of Z_t as functions of Ψ. Specifically,

$$E\left[X\left(Z_t\right)\right] = f\left(\Psi\right) \tag{104}$$

where $f(\cdot)$ is also a vector valued function (appendices provide the exact formulae).

This immediately suggests a simple GMM based estimation strategy. The GMM moment conditions are,

$$g_{1T}\left(Z_t;\Psi_0\right) = \frac{1}{T}\sum_{t=1}^{T} X\left(Z_t\right) - f\left(\Psi_0\right). \tag{105}$$

Moreover, the additive separability of data and parameters in Equation (105) suggests a 'fixed' optimal GMM weighting matrix free from any particular parameter vector and based on the data alone. We denote the data used as $\widetilde{X}_T = \{X_1, X_2..., X_T\}$. The optimal GMM weighting matrix is the inverse of the spectral density at frequency zero of $g_{1T}\left(\widetilde{X}_T;\Psi_0\right)$, which we denote as $S_{11}\left(\widetilde{X}_T\right)$, because only the first term on the right hand side of Equation (105) contains any random variables (data).

Further, to reduce the number of parameters implicitly estimated in calculating the optimal GMM weighting matrix while still accommodating high persistence in the orthogonality conditions, we exploit the structure implied by the model. In particular, we compute the spectral density in two steps. First, we consider the spectral density of $Y_t^c = \left[Y_t^P, Y_{t-1}^P\right]'$, where Y_t^P is an observable proxy for the state vector (which includes a latent variable, q_t). In practice, we use $Y_t^P = [\Delta c_t, u_t, \pi_t, rf_t]'$with rf_t proxying for q_t. Because Y_t^P is quite persistent, we use a standard VAR(1) pre-whitening technique. Denote the spectral density at frequency zero estimate of Y_t^c as $\widehat{S_{11}}\left(Y_T^c\right)$. Second, we project $\widetilde{X}_T$ onto Y_t^c. Let $\widehat{B}$ denote the least squares projection coefficients and $\widehat{D}$ the (diagonal) variance-covariance matrix of the residuals of this projection. Then, our estimate for $\widehat{S_{11}}(Z_T)$ is

$$\widehat{S_{11}}\left(\widetilde{X}_T\right) = \widehat{B}\widehat{S_{11}}\left(Y_T^c\right)\widehat{B}' + \widehat{D} \tag{106}$$

The inverse of $\widehat{S_{11}}\left(\widetilde{X}_T\right)$ is the optimal weighting matrix. To estimate the system, we minimize the standard GMM objective function,

$$J\left(W_T;\widehat{\Psi}\right) = g_{1T}\left(\widehat{\Psi}\right)\cdot\widehat{S_{11}}\left(\widetilde{X}_T\right)^{-1}\cdot g_{1T}\left(\widehat{\Psi}\right)' \tag{107}$$

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.

The parameters from this stage are then used as starting values for the conditional estimation described in the main text.