Asset Pricing with Regime Shifts in Consumption and Dividend Growth[*]

JOB MARKET PAPER

Anisha Ghosh[†]

this version: November 20, 2008

Abstract

We present an asset pricing model with two regimes, where the conditional mean of the aggregate consumption growth rate reverts to its unconditional mean with a process that differs across regimes. The representative consumer, endowed with recursive preferences, observes the conditional mean but not the regime. He forms a posterior probability of being in the first regime based on his current information set. The two state variables, the conditional mean and the posterior probability, are latent to the econometrician. We show that, in equilibrium, the market-wide log price-dividend ratio and risk free rate and, hence, the pricing kernel are affine functions of the two state variables and their product. The two equations may, therefore, be inverted to extract the two latent state variables as known functions of the observable price-dividend ratio and risk free rate. The two-regime model produces significantly lower pricing errors in the market portfolio, the risk free rate, and the cross-section of returns over 1930-2006 than the single-regime model does. It also performs significantly better at predicting market returns, the equity premium, and the cross-section of returns compared to linear predictive regressions of returns on the lagged price-dividend ratio and interest rate. The first regime, characterized by a short (less than 3 years) half-life and higher volatility of the conditional mean of the consumption growth rate, is about twice as likely to correspond to NBER-registered recessions than the second regime characterized by a half-life longer than 8 years.

---

[*] I thank Alan Bester, Stephane Guibaud, Christian Julliard, Oliver Linton, Alex Michaelides, Pietro Veronesi, seminar participants at LSE, and especially George Constantinides for helpful comments. I remain responsible for errors and omissions.

[†] Department of Economics, London School of Economics, e-mail: a.ghosh@lse.ac.uk

1 Introduction

The Great Depression and its slow recovery in the thirties, World War II, the "irrational exuberance" of the nineties, the dotcom crash at the turn of the century, and the current credit crisis are suggestive of regime shifts in the economy. In this paper, we explore the extent to which regime shifts in the mean reversion and variance of the conditional mean of the aggregate consumption and dividend growth rates improve our understanding of the level of capital asset prices and the cross-section of their returns.[1]

We identify two distinctly different patterns in the mean reversion of the conditional mean of the aggregate consumption and aggregate dividend growth rates, in a two-regime model. In the first regime, the half-life of the conditional mean is shorter than three years, while in the second regime, it is longer than three years. The volatility of the conditional mean in the first regime is about double of that in the second regime. More to the point, the two-regime model produces significantly lower pricing errors in the market portfolio, the risk free rate, and the cross-section of asset returns over the entire available sample period 1930-2006 than the single-regime model does. Furthermore, the over-identifying restrictions are not rejected, even though they are strongly rejected in the single-regime model.

Several extensions of the standard neoclassical model of capital asset prices address the failure of the standard model to explain the observed behavior of the equity premium, the risk free rate, and the cross-section of capital asset returns.[2] One line of research recognizes that reversion of the conditional means of the aggregate consumption and dividend growth rates to their unconditional means is slow. In conjunction with Kreps and Porteus (1978) preferences, the slow reversion of the conditional means holds promise in better understanding the puzzles.[3] Evidence on the success of these

---

[1] Pastor and Stambaugh (2001) document evidence of breaks in the equity premium in the early thirties and forties, and in the early and mid-nineties. Koijen and Van Binsbergen (2007) document both a transient and a persistent component in the conditional mean of the aggregate dividend growth rate. Lettau, Ludvigson, and Wachter (2008) find evidence of a break in the consumption volatility around 1992 followed by a break in the log price-dividend ratio of the market around 1995. Lettau and Van Nieuwerburgh (2008) report evidence of two breaks in the mean of the aggregate price-dividend ratio around 1954 and 1994. An extensive body of work in the macroeconomic literature that finds evidence of a regime shift to lower volatility of real macroeconomic activity occurring in the last 15 years of the 20th century (Kim and Nelson (1999), McConnell and Perez-Quiros (2000), Blanchard and Simon (2001), Stock and Watson (2002), Lettau, Ludvigson and Wachter (2008)).

[2] This extensive literature is reviewed in a collection of essays in Mehra (2008); the textbooks by Campbell, Lo, and MacKinlay (1997) and Cochrane (2005); and the articles by Cochrane and Hansen (1992), Kocherlakota (1996), Campbell (2000, 2003), Constantinides (2002), and Mehra and Prescott (2003).

[3] See Bansal and Yaron (2004), Alvarez and Jerman (2005), Bansal, Dittmar, and Lundblad (2005), Bekaert, Engstrom, and Xing (2005), Kiku (2006), Malloy, Moskowitz, and Vissing-Jorgensen (2006), Bansal, Gallant, and Tauchen (2007), Bansal, Kiku, and Yaron (2007), Hansen and Scheinkman

models is mixed.[4] Furthermore, Constantinides and Ghosh (2008) report that the point estimate of the mean reversion in a single-regime model is unstable in subperiods, with the half-life of the conditional mean being shorter than three years in some cases and longer than three years in others. The latter observation motivates the approach of this paper that recognizes a shift in regimes.

We introduce an economy with two regimes. The mean reversion and variance of the conditional mean of the aggregate consumption growth rate differ across regimes. Likewise, the mean reversion and variance of the conditional mean of the aggregate dividend growth rate differ across regimes. The representative consumer observes the consumption and dividend levels, and the conditional means of their growth rates, but not the regime. The consumer forms a posterior probability of being in the first regime, based on the history of the observed consumption and dividend levels, the conditional means of their growth rates, and other information which is generally unavailable to the econometrician. Thus the conditional means of the growth rates and the posterior probability of being in the first regime are potentially latent state variables to the econometrician. The consumer is assumed to be endowed with the version of Kreps and Porteus (1978) preferences adopted by Epstein and Zin (1989) and Weil (1989). A nice feature of our model is that the latent state variables become observable to the econometrician because, in the log-linearized version of our model, the aggregate log price-dividend ratio and the risk free rate are affine functions of the state variables and their product. Hence, we can invert the system to express the two state variables as known functions of the observable aggregate log price-dividend ratio and interest rate.

We examine the empirical plausibility of the model using the Simulated Moments Estimation (SMM) approach of Duffie and Singleton (1993). The stochastic discount factor in the economy is a function of the two state variables and their cross product. Since the market-wide log price-dividend ratio and the risk free rate are affine functions of the state variables and their product, we rewrite the pricing kernel as a function of the observable log price-dividend ratio and the risk free rate. We substitute this expression for the pricing kernel into the set of Euler equations to obtain a set of moment restrictions that are expressed entirely in terms of observables. Also, the time-series specification of the model imposes constraints on the unconditional moments of the aggregate consumption and dividend growth rates. These are additional constraints,

---

(2007), Hansen, Heaton, and Li (2008), and Lettau and Ludvigson (2008) .

[4] Bansal and Yaron (2004), Kiku (2006), Malloy, Moskowitz, and Vissing-Jorgensen (2006), Bansal, Gallant, and Tauchen (2007), and Bansal, Kiku, and Yaron (2007) report positive evidence. Constantinides and Ghosh (2008) and Lettau and Ludvigson (2008) report negative evidence. The negative evidence in Constantinides and Ghosh (2008) is due, in part, to the empirical methodology of the paper which applies powerful tests based on the observation that potentially latent state variables are observable because both the aggregate price-dividend ratio and the interest rate are functions only of these state variables under the null; and the added restriction that the unconditional moments of consumption growth and aggregate dividend growth impose constraints in addition to the pricing constraints.

in addition to the pricing restrictions.

We first consider the plausibility of the time-series specification of the model. Using aggregate consumption and dividend data alone, and the constraints imposed on the unconditional moments of the aggregate consumption and dividend growth rates by the model, we identify two distinctly different patterns in the mean reversion of the conditional mean of the aggregate consumption and aggregate dividend growth rates. In the first regime, the half-life of the conditional mean is shorter than three years, while in the second regime, it is longer than three years. The volatility of the conditional mean in the first regime is about double of that in the second regime. Moreover, the over-identifying restrictions test fails to reject the time-series specification of the model.

Next we test the ability of the model to jointly explain the pricing restrictions, given by the Euler equations for a set of assets, and the restrictions on the unconditional moments of aggregate consumption and dividend growth rates implied by the time-series specification of the model. We find that the two-regime model produces significantly lower pricing errors in the market portfolio, the risk free rate, and the cross-section of asset returns than the single-regime model does and that the over-identifying restrictions are not rejected.

The time-series of the extracted state variables reveal that the first regime, characterized by a short (less than 3 years) half-life and higher volatility of the conditional mean of aggregate consumption and dividend growth rates, is more likely to correspond to NBER registered recession periods. The second regime, characterized by a half-life of more than 3 years and lower volatility of the conditional mean, is more likely to correspond to the expansionary phase of the business cycle.

Finally, we examine the ability of the regime shifts model to predict stock market returns and the equity premium. The model implies analytical expressions for the return on the market portfolio and the equity premium in terms of the two state variables. Using the time-series of the extracted state variables and performing in-sample least squares regressions of the form implied by the model leads to an *adjusted*-R^2 of 7.3% for the market return and 10.3% for the equity premium for the full sample period 1930-2006. These are much larger than those obtained from linear predictive regressions of the market returns and the equity premium on the lagged aggregate price-dividend ratio and the risk free rate (3.9% and 7.3%, respectively). The *adjusted*-R^2 from the model-implied regressions rises to 15.0% and 19.6% for the market returns and the equity premium, respectively, when attention is restricted to the post war subperiod 1947-2006. As for the full sample, these are also much larger than those obtained from linear predictive regressions of the market returns and the equity premium on the lagged aggregate price-dividend ratio and the risk free rate (8.0% and 12.0%, respectively). The model also has superior out-of-sample predictive performance for the market returns and the equity premium and in-sample predictive performance for the cross-section of returns compared to linear predictive regressions of returns on the lagged aggregate price-dividend ratio and the risk free rate. These findings lend additional support to the regime shifts model highlighted in the paper.

The paper is organized as follows. In Section 2, we present the two-regime model. In Section 3, we describe our estimation methodology. The data are discussed in Section 4. Section 5 presents the empirical evidence on the time-series specification of the model, the ability of the model to explain the risk free rate, the equity premium, and the cross-section of asset returns, and the economic interpretation of the two regimes identified by the data. Section 6 examines the ability of the model to predict stock market returns, the equity premium, and the cross-section of returns. Section 7 concludes. The appendix contains derivation of the main theoretical results and details of the estimation methodology.

2 The Model

We consider the following specification of the time-series processes of aggregate consumption and dividend growth rates, which is an extension of the single-regime model of Bansal and Yaron (2004):

$$\begin{aligned} x_{t+1} &= \rho_{s_{t+1}} x_t + \varphi_e \sigma_{s_{t+1}} e_{t+1}, \\ \Delta c_{t+1} &= \mu + x_t + \sigma_{s_{t+1}} \eta_{t+1}, \\ \Delta d_{t+1} &= \mu_d + \phi x_t + \varphi_d \sigma_{s_{t+1}} u_{t+1}, \end{aligned} \tag{1}$$

where c_{t+1} is the logarithm of the aggregate consumption level; d_{t+1} is the logarithm of the aggregate stock market dividends; x_t is the state variable that simultaneously drives the conditional means of aggregate consumption and dividend growth rates; and $s_t = 0, 1$ is a second state variable that denotes the economic regime. The persistence parameter of the conditional mean of consumption and dividend growth, ρ_{s_t}, and the level of its volatility, σ_{s_t}, are generally different in the two regimes. We assume $0 < \rho_0 < \rho_1 < 1$, and $0 < \sigma_0$, $0 < \sigma_1$.

The representative consumer in the economy observes the conditional mean, x_t, but not the regime, s_t.[5] The consumer's posterior probability at time t of being in regime $s_t = 0$, given his information set, $\mathcal{F}(t)$, is

$$p_t \equiv Prob\left(s_t = 0 | \mathcal{F}(t)\right) \tag{2}$$

The information set $\mathcal{F}(t)$ includes the history of consumption, dividends, the conditional mean of consumption growth, and any other information that the consumer uses

---

[5] Note that, for the time-series specification (1), knowledge of x_t is not sufficient to perfectly infer ρ_{s_t} or σ_{s_t}, and hence, the regime s_t.

to form expectations. The econometrician's information set is generally a weak subset of $\mathcal{F}(t)$ and, hence, the probability, p_t, and the conditional mean of consumption growth, x_t, are latent state variables.

We assume that s_t follows a Markov process with the following transition probability matrix:

$$\Pi = \begin{pmatrix} \pi_0 & 1-\pi_1 \\ 1-\pi_0 & \pi_1 \end{pmatrix}, \tag{3}$$

where $0 < \pi_i < 1$ for $i = 0, 1$. Thus, the consumer's probability of being in regime $s_{t+1} = 0$ at time $t+1$, given his information set, $\mathcal{F}(t)$, is

$$Prob\left(s_{t+1} = 0|\mathcal{F}(t)\right) = \pi_0 p_t + (1-\pi_1)(1-p_t) \equiv f(p_t) \tag{4}$$

Note that $0 < f(p_t) < 1$ for all p_t, $0 \leq p_t \leq 1$.

Once the consumer updates his information set at time $t+1$, his probability of being in regime $s_{t+1} = 0$ at time $t+1$ is $p_{t+1} \equiv \Pr ob\left(s_{t+1} = 0|\mathcal{F}(t+1)\right)$. We assume that the consumer's expectations are rational in that

$$E\left[p_{t+1}|\mathcal{F}(t)\right] = Prob\left(s_{t+1} = 0|\mathcal{F}(t)\right). \tag{5}$$

In other words, we have

$$p_{t+1} = f(p_t) + \varepsilon_{t+1}, \tag{6}$$

where $E\left[\varepsilon_{t+1}|\mathcal{F}(t)\right] = 0$, i.e. ε_{t+1} is a martingale difference sequence and is uncorrelated with all past information $\mathcal{F}(t)$. It is easily shown that the unconditional mean of p_t is $\frac{1-\pi_1}{2-\pi_0-\pi_1}$ and the unconditional mean of x_t is zero. Also, given π_i, $i = 0, 1$, the density of the duration d_i of regime i is $f(d_i|\pi_i) = \pi_i^{d_i-1}(1-\pi_i)$, and its mean is $E(d_i|\pi_i) = \frac{1}{1-\pi_i}$. Further, we assume that ε_{t+1} is uniformly distributed over $[-a(p_t),\ a(p_t)]$ where $a(p_t) \equiv min\left[f(p_t), 1-f(p_t)\right]$. Then, it is easily shown that $0 \leq p_{t+1} \leq 1$ for all p_t, $0 \leq p_t \leq 1$.[6] Finally, the time-series specification is assumed to hold at the annual frequency.

---

[6] The solution of the model for asset prices and the derivation of the resulting Euler equations does not require us to specify any distributional assumptions for ε_{t+1} in equation (6) other than its martingale difference property. In our empirical analysis, we examine the ability of the model to simultaneously explain the pricing restrictions given by the Euler equations and the restrictions on the unconditional moments of aggregate consumption and dividend growth rates implied by the time-series specification of the model. The time-series specification of the model in equations (1), (3), and (6) implies that the moments of consumption and dividend growth rates do not have analytic representations in terms of the unknown time-series parameters. Hence, to estimate the parameters and perform an overidentifying restrictions test to examine the empirical validity of the model specification, we rely on the Simulated Moments Estimation (SME) approach of Duffie and Singleton (1993). This requires us to specify the distribution of ε_{t+1} in equation (6).

We assume that the consumer has the version of Kreps and Porteus (1978) preferences adopted by Epstein and Zin (1989) and Weil (1989). These preferences allow for a separation between the coefficient of risk aversion and the elasticity of intertemporal substitution. The utility function is defined recursively as

$$V_t = \left[(1-\delta)C_t^{\frac{1-\gamma}{\theta}} + \delta\left(E\left[V_{t+1}^{1-\gamma}|\digamma(t)\right]\right)^{\frac{1}{\theta}}\right]^{\frac{\theta}{1-\gamma}} \tag{7}$$

where δ denotes the subjective discount factor, $\gamma > 0$ is the coefficient of risk aversion, $\theta = \frac{1-\gamma}{1-\frac{1}{\psi}}$, and $\psi > 0$ is the elasticity of intertemporal substitution. Note that the sign of θ depends on the relative magnitudes of γ and ψ. The standard time-separable power utility is obtained as a special case when $\theta = 1$, i.e. $\gamma = \frac{1}{\psi}$.

For this specification of preferences, Epstein and Zin (1989) and Weil (1989) show that, for any asset j, the first-order conditions of the consumer's utility maximization yield the following Euler equations,

$$E\left[\exp(m_{t+1} + r_{j,t+1})|\digamma(t)\right] = 1, \tag{8}$$

$$m_{t+1} = \theta \log \delta - \frac{\theta}{\psi}\Delta c_{t+1} + (\theta - 1)r_{c,t+1}, \tag{9}$$

where m_{t+1} is the natural logarithm of the intertemporal marginal rate of substitution, $r_{j,t+1}$ is the continuously compounded return on asset j, and $r_{c,t+1}$ is the unobservable continuously compounded return on an asset that delivers aggregate consumption as its dividend each period.

We rely on log-linear approximations for the log return on the consumption claim, $r_{c,t+1}$, and that on the market portfolio (the return on the aggregate dividend claim), $r_{m,t+1}$, as in Campbell and Shiller (1988),

$$\begin{aligned} r_{c,t+1} &= \kappa_0 + \kappa_1 z_{t+1} - z_t + \Delta c_{t+1}, & (10) \\ r_{m,t+1} &= \kappa_{0,m} + \kappa_{1,m} z_{m,t+1} - z_{m,t} + \Delta d_{t+1}, & (11) \end{aligned}$$

where z_t is the log price-consumption ratio and $z_{m,t}$ the log price-dividend ratio. In equation (10), $\kappa_1 = \frac{e^{\overline{z}}}{1+e^{\overline{z}}}$ and $\kappa_0 = log(1+e^{\overline{z}}) - \kappa_1\overline{z}$ are log-linearization constants, where $\overline{z}$ denotes the long run mean of the log price-consumption ratio. Similarly, in equation (11), $\kappa_{1,m} = \frac{e^{\overline{z}m}}{1+e^{\overline{z}m}}$ and $\kappa_{0,m} = log(1+e^{\overline{z}_m}) - \kappa_1\overline{z}_m$, where $\overline{z}_m$ denotes the long run mean of the log price-dividend ratio.

Note that the current model specification involves two state variables, x_t and p_t. We conjecture and verify the following expressions for the log price-consumption ratio and log price-dividend ratio at date t, respectively, (see Appendices A.1 and A.2 for derivations and expressions for the parameters $A_0(0)$, $A_1(0)$, $A_0(1)$, $A_1(1)$, $A_{0,m}(0)$, $A_{1,m}(0)$, $A_{0,m}(1)$, and $A_{1,m}(1)$):

$$\begin{aligned} z_t &= p_t\left[A_0(0)+A_1(0)x_t\right]+(1-p_t)\left[A_0(1)+A_1(1)x_t\right], & (12)\\ z_{m,t} &= p_t\left[A_{0,m}(0)+A_{1,m}(0)x_t\right]+(1-p_t)\left[A_{0,m}(1)+A_{1,m}(1)x_t\right]. & (13)\end{aligned}$$

The intuition for the above result is straightforward. The log price-consumption ratio in equation (12) takes the value $A_0(0)+A_1(0)x_t$, if the consumer knows with certainty that the economy is in regime 0 at date t and takes the value $A_0(1)+A_1(1)x_t$, if the consumer knows with certainty that the economy is in regime 1. The uncertainty about the current regime leads him to price the claim as the sum of the valuation in each regime, weighted by the posterior probability of each regime. A similar intuition applies to equation (13).

The gross risk free rate, $R_{f,t}$, between periods t and $t+1$, is a function of the two latent state variables and their product (see Appendix A.3 for derivation and expressions for the parameters $A_{0,f}$, $A_{1,f}$, $A_{2,f}$, $A_{3,f}$),

$$\frac{1}{R_{f,t}}=A_{0,f}+A_{1,f}x_t+A_{2,f}p_t+A_{3,f}p_tx_t. \tag{14}$$

Equations (11), (13), and (14) imply that the equity premium is given by the expression (see Appendix A.4 for derivations and expressions for the parameters E_0, E_1, E_2, E_3, and E_4)

$$E\left[\left(r_{m,t+1}-\frac{1}{R_{f,t}}\right)|\digamma(t)\right]=E_0+E_1x_t+E_2p_t+E_3p_tx_t+E_4p_t^2x_t. \tag{15}$$

Note that the Bansal and Yaron (2004) model without stochastic volatility is obtained as a special case when $\pi_0=\pi_1=1$, i.e. each of the two states is an absorbing one. In the absense of stochastic volatility, the model makes the counterfactual prediction of a constant equity premium. To obtain a time-varying premium, Bansal and Yaron (2004) introduce a stochastically evolving variance which is assumed to follow an AR(1) process with the consequence that the process can take negative values with non-zero probability. In contrast, the specification of the time-series processes in equation (1) ensures that the variance never takes negative values while at the same time predicting time variation in the equity premium. Moreover, as revealed by equation (15), the equity premium is a non-linear function of the state variables, x_t and p_t.

3 Estimation Methodology

We evaluate the model using the Euler equations for a set of logarithmic portfolio returns as well as the restrictions on the unconditional moments of aggregate consumption and dividend growth rates implied by the time-series specification of the model. The Euler equations in (8) and (9) are restated here for convenience:

$$E_t\left[\exp(m_{t+1}+r_{j,t+1})\right]=1,$$

where m_{t+1} is the pricing kernel,

$$m_{t+1}=\theta\log\delta-\frac{\theta}{\psi}\Delta c_{t+1}+(\theta-1)r_{c,t+1}.$$

Substituting the log-affine approximation for $r_{c,t+1}$ (equation (10)) into the expression for the pricing kernel (equation (9)), and noting that z_t is given by equation (12), we have, (see Appendix A.5 for details),

$$m_{t+1}=c_0+c_1\Delta c_{t+1}+c_2p_{t+1}+c_3p_t+c_4x_{t+1}+c_5x_t+c_6p_{t+1}x_{t+1}+c_7p_tx_t, \tag{16}$$

where,

$$\begin{aligned}
c_0 &= \theta\log(\delta)+(\theta-1)\kappa_0+(\theta-1)(\kappa_1-1)A_0(1),\\
c_1 &= -\frac{\theta}{\psi}+\theta-1,\\
c_2 &= (\theta-1)\kappa_1\left[A_0(0)-A_0(1)\right],\\
c_3 &= -(\theta-1)\left[A_0(0)-A_0(1)\right]\\
c_4 &= (\theta-1)\kappa_1A_1(1),\\
c_5 &= -(\theta-1)A_1(1),\\
c_6 &= (\theta-1)\kappa_1\left[A_1(0)-A_1(1)\right],\\
c_7 &= -(\theta-1)\left[A_1(0)-A_1(1)\right].
\end{aligned}$$

The above specification of the pricing kernel involves the latent state variables p_t and x_t. We proceed to rewrite the pricing kernel in terms of observables alone. We note that the market-wide log price-dividend ratio, given by equation (13), is a function only of the two latent state variables,

$$z_{m,t}=p_t\left[A_{0,m}(0)+A_{1,m}(0)x_t\right]+(1-p_t)\left[A_{0,m}(1)+A_{1,m}(1)x_t\right].$$

Also, the gross risk free rate is a function only of these two state variables (see equation (14)),

$$\frac{1}{R_{f,t}}=A_{0,f}+A_{1,f}x_t+A_{2,f}p_t+A_{3,f}p_tx_t.$$

The pair of nonlinear equations, (13) and (14), may be inverted to express the latent state variables, x_t and p_t, as known functions of the observables, $z_{m,t}$ and $R_{f,t}$ (see Appendix A.5 for details). Substituting these expressions for x_t and p_t into the pricing kernel, equation (16), we have an expression for the pricing kernel entirely in terms of observables. The specification of preferences and the time-series processes imposes restrictions on the parameters of the pricing kernel. We substitute this expression for the pricing kernel into the set of Euler equations (8) to obtain a set of moment restrictions that are expressed entirely in terms of observables.

We estimate and test the model using annual data over the entire available sample period 1930-2006. We first test the hypothesized time-series processes, including the Markov process. We do this by estimating the time-series parameters and performing an over-identifying restrictions test using restrictions on the unconditional moments of aggregate consumption and dividend growth rates implied by the time-series specification of the model (equation (1)). In particular, we include moments corresponding to the unconditional means, variances, and first, second, and third-order autocovariances of consumption and dividend growth rates, the covariance between consumption and dividend growth rates, and the covariance between consumption growth and one and two lags of the dividend growth rate. This gives 13 moment restrictions corresponding to the time-series specification of the model. The number of time-series parameters to be estimated is 11. Since the moments do not have a closed-form representation, we rely on the Simulated Moments Estimation approach of Duffie and Singleton (1993) (see Appendix A.6 for details of the approach).

We next examine the ability of the model to simultaneously explain the pricing restrictions given by the Euler equations and the time-series moment restrictions implied by the time-series specification of the model. We first consider the case when the asset menu consists of the market portfolio and the risk free rate. The lagged log price-dividend ratio of the market and the lagged log risk free rate are used as instruments. The Euler equations for the two assets along with the two chosen instruments give 6 moment restrictions. To this set of pricing restrictions, we add the 13 moment restrictions implied by the time-series specification of the model in equations (1), (3), and (6). Thus, we have a total of 19 moment conditions. The total number of parameters to be estimated is 14, including 11 time-series parameters and 3 preference parameters. We estimate the parameters with the SME approach and test the specification of the model using the overidentifying restrictions.[7]

Finally, we examine the ability of the model to explain the cross-section of returns.

---

[7] See Appendix A.6 for an extension of the methodology in Duffie and Singleton (1993) that accommodates the possibility that a subset of the moment restrictions have analytic representations in terms of observable variables and the unknown parameter vector while the others do not. We use this extension to compute the standard errors of the parameter estimates and obtain the test statistic for over-identifying restrictions since the Euler equations deliver moment restrictions that are known functions of observables like the aggregate consumption growth rate, the market-wide price-dividend ratio, the risk free rate, and the returns on the market portfolio and the risk free rate, and the unknown time-series and preference parameters, but the time-series moment restrictions including the means, variances and autocovariance functions of aggregate consumption and dividend growth rates do not have analytic representations.

In this case, the asset menu consists of the market portfolio, the risk free rate, and portfolios of "Small" capitalization, "Large" capitalization, "Growth" and "Value" stocks. The Euler equations for the 6 assets give 6 moment restrictions. To this set of pricing restrictions, we add the 13 moment restrictions implied by the time-series specification of the model. This gives, once again, a total of 19 moment conditions in 14 parameters. We estimate the parameters and test the model specification with the SME approach.

We note that the size of the historical sample is small (77) relative to the number of parameters to be estimated (11 time-series parameters and 3 preference parameters) and the moment restrictions are highly non-linear functions of the parameters. Therefore, we base inference of the parameters and testing of the overidentifying restrictions on both, the asymptotic properties of the SME methodology and the finite-sample distributions obtained via Monte-Carlo simulations. The details of the simulation design are described in Appendix A.7.

4 Data

We estimate the model at the annual frequency, using annual data over the entire available sample period 1930-2006. The asset menu consists of the market, the risk free rate, and portfolios of "Value", "Growth", "Small" capitalization, and "Large" capitalization stocks. Our market proxy is the Centre for Research in Security Prices (CRSP) value-weighted index of all stocks on the NYSE, AMEX, and NASDAQ. The proxy for the risk free rate is the one-month Treasury Bill rate (from Ibbotson Associates). The construction of the size and book-to-market portfolios is as in Fama and French (1993). In particular, for the size sort, all NYSE, AMEX, and NASDAQ stocks are allocated across 10 portfolios according to their market capitalization at the end of June of each year. Value-weighted returns on these portfolios are then computed over the following twelve months. NYSE breakpoints are used in the sort. "Small" and "Large" denote the bottom and top market capitalization deciles, respectively. Similarly, value-weighted returns are computed for portfolios formed on the basis of BE/ME at the end of June of each year using NYSE breakpoints. The BE used in June of year t is the book equity for the last fiscal year end in $t-1$ and ME is the price times shares outstanding at the end of December of $t-1$. "Growth" and "Value" denote the bottom and top BE/ME deciles, respectively. Annual returns for the above portfolios are computed by compounding monthly returns within each year. Also used in the empirical analysis are the price-dividend ratios and dividend growth rates of the above mentioned portfolios. Data on these are obtained from the CRSP files. All nominal quantities are converted to real, using the personal consumption deflator.

Table 1 provides descriptive statistics for the continuously compounded returns, the log price-dividend ratios, and the log dividend growth rates for the six assets mentioned above, for the annual sample over the period 1930-2006. The table illustrates the well documented equity premium and the size and value premia. Over the sample period, the annual equity premium over the 1-month Treasury bill rate has mean 5.8% and volatility of market returns is 19.3%. The annual risk free rate has mean 0.8% and standard deviation 5.0%. The annual mean premium of small over large stocks is 4.5% and of value over growth stocks is 4.1%. Value stocks are much more volatile than growth stocks and small stocks are much more volatile than large stocks.

The annual log price-dividend ratio on the market has a mean of 3.27 and standard error of 0.38 over the sample period. The price-dividend ratios of the "Small" and "Value" portfolios are much more volatile with annual volatilities at 0.71 and 1.14, respectively, compared to their counterparts, namely the "Large" and "Growth" portfolios that have volatilities 0.44 and 0.63, respectively.

The average annual log dividend growth rate on the market portfolio is 1.4% with volatility 10.8%. The mean and volatility of the "Small" (8.3% and 34.7%) and "Value" (7.0% and 56.8%) portfolios are much higher compared to their counterparts, namely the "Large" (1.2% and 13.6%) and "Growth" (0.7% and 20.6%) portfolios.

Finally, for consumption, we use real per capita consumption of non-durables and services from the National Income and Product Accounts (NIPA). We make the standard "end-of-period" timing assumption that consumption during period t takes place at the end of the period. Growth rates are constructed by taking the first difference of the corresponding log series. The annual log consumption growth has a mean of 1.5% and standard deviation of 2.6% over the sample period.

5 Empirical Results

5.1 Empirical Evidence on the Time-Series Specification

We test the hypothesized time-series processes, including the Markov process. We do this by estimating the time-series parameters and performing an over-identifying restrictions test using the restrictions on the unconditional moments of consumption and dividend growth rates implied by the time-series specification of the model. In particular, we include the following 13 moments of aggregate consumption and dividend growth rates: $E(\Delta c_t)$, $Var(\Delta c_t)$, $Cov(\Delta c_t, \Delta c_{t+1})$, $Cov(\Delta c_t, \Delta c_{t+2})$, $Cov(\Delta c_t, \Delta c_{t+3})$, $E(\Delta d_t)$, $Var(\Delta d_t)$, $Cov(\Delta d_t, \Delta d_{t+1})$, $Cov(\Delta d_t, \Delta d_{t+2})$, $Cov(\Delta d_t, \Delta d_{t+3})$, $Cov(\Delta c_t, \Delta d_t)$, $Cov(\Delta c_{t+1}, \Delta d_t)$, and $Cov(\Delta c_{t+2}, \Delta d_t)$. The number of time-series parameters to be estimated is 11: $\beta = (\mu, \mu_d, \phi, \varphi_d, \rho_0, \rho_1, \varphi_e, \sigma_0, \sigma_1, \pi_0, \pi_1)'$ with true value β_0. This gives us 2 over-identifying restrictions which can be used to test the specification of the model. Now, the time-series specification of the model in equations (1), (3), and (6) implies that the unconditional moments of consumption and dividend growth do not have analytic representations in terms of the unknown parameter vector. Hence, we rely on the Simulated Moments Estimation approach of Duffie and Singleton (1993).

The *simulated moments estimator* (SME) is a value of β chosen to minimize the distance between the sample mean computed from simulated data and the historical sample mean (see Appendix A.6 for details),

$$\widehat{\beta}_{SME} = \underset{\beta \in \Theta}{arg\ min} \left\{ G_T(\beta)' W_T G_T(\beta) \right\}$$

where $G_T(\beta) = \left[\frac{1}{T} \sum_{t=1}^{T} f_t^* - \frac{1}{N} \sum_{s=1}^{N(T)} f_s^\beta \right]$, $\{f_t^*\}_{t=1}^{T}$ denotes the historical observations, $\left\{ f_t^\beta \right\}_{t=1}^{N(T)}$ denotes the observations from the simulated sample, T is the size of the historical sample, $N(T)$ denotes the size of the simulated sample, and W_T is a sequence of positive semi-definite weighting matrices.

Note that, for the time-series moment restrictions considered in this section, f_t^* does not depend on the parameter vector β. In this case, under the regularity conditions in Duffie and Singleton (1993), if the weighting matrix W_T is chosen such that $W_T \rightarrow W_0 = \Sigma_0^{-1}$ almost surely, where (for any t)

$$\Sigma_0 \equiv \sum_{j=-\infty}^{\infty} E \left(\left[f_t^* - E\left(f_t^*\right) \right] \left[f_{t-j}^* - E\left(f_{t-j}^*\right) \right]' \right), \tag{17}$$

then $\sqrt{T} \left(\widehat{\beta}_{SME} - \beta_0 \right)$ converges in distribution as $T \rightarrow \infty$ to a normal random vector with mean zero and covariance matrix

$$\Lambda^{eff} = (1 + \tau) \left(D_0' \Sigma_0^{-1} D_0 \right)^{-1},$$

where $D_0 = E\left(\frac{\partial}{\partial \beta'} f_\infty^{\beta_0} \right)$ and $\frac{T}{N(T)} \rightarrow \tau$ as $T \rightarrow \infty$.

As with the GMM approach, an over-identifying restrictions test may be performed to test the specification of the model. Given the normalized asymptotic distribution of the estimator, the J-stat converges in distribution as $T \rightarrow \infty$ to a chisquared random variable with $q-p$ degrees of freedom under the null that the model is correctly specified (q denotes the number of moment restrictions and p the number of parameters to be estimated):

$$J \equiv T \left\{ G_T(\widehat{\beta}_{SME})' \left[(1+\tau)' \, \widehat{\Sigma} \right]^{-1} G_T(\widehat{\beta}_{SME}) \right\} \xrightarrow{d} \chi_{q-p}^2,$$

where $\widehat{\Sigma}$ denotes a consistent estimator of Σ_0.

For any positive definite weighting matrix W_T other than the efficient weighting matrix Σ_0^{-1}, $\sqrt{T}\left(\widehat{\beta}_{SME}-\beta_0\right)$ converges in distribution as $T \to \infty$ to a normal random vector with mean zero and covariance matrix,

$$\Lambda^{non-eff} = (1+\tau)\left(D_0' W_T D_0\right)^{-1} D_0' W_T \Sigma_0 W_T D_0 \left(D_0' W_T D_0\right)^{-1}. \qquad (18)$$

The "Dist" statistic (the analogue of the J-stat for any positive definite weighting matrix other than the efficient weighting matrix), $Dist \equiv T\left\{G_T(\widehat{\beta}_{SME})' W_T G_T(\widehat{\beta}_{SME})\right\}$, has a nonstandard asymptotic distribution. However, the p-values of the computed statistic can still be consistently computed to test the null hypothesis that the model is correctly specified (see Jagannathan and Wang (1996) and Parker and Julliard (2005) for derivations of the asymptotic distribution and computation of the p-values when the moment conditions are linear and nonlinear, respectively, in the parameters).

We perform our estimation and tests using both the identity and the efficient weighting matrices. We use annual data on aggregate consumption and dividend growth rates over the period 1930-2006. Hence, the size of the historical sample, T, is 77. We choose the simulation sample size, N, to be 5000. This choice ensures that $\tau \approx 0$ and, hence, that the SME estimators are asymptotically almost as efficient as when the functional forms of the population moments were known. We estimate Σ_0 using the Newey-West estimator with two lags.

The estimation results are reported in Table 2. Panels A and B report results for the identity and the efficient weighting matrices, respectively. The first row of Panel A reports the point estimates of the time-series parameters along with the associated standard errors in parentheses. The persistence parameter of the conditional mean of consumption growth in the two regimes takes values 0.92 and 0.50, respectively. This suggests that in regime 0, the conditional mean of consumption and dividend growth has a half-life of just over 8 years, i.e. it has a frequency much longer than the average period of the business cycle. However, in regime 1, the consumption and dividend dynamics are driven by a much higher frequency component that has a half-life of just 1 year. The volatility of the conditional mean takes values 1.2% and 2.1%, respectively, in the two regimes. These results suggest the presence of two regimes, one in which consumption and dividend growth rates are more persistent and less volatile and the other during which the growth rates are much less persistent and have higher volatility. This is consistent with the findings in Koijen and Van Binsbergen (2007) who document both a transient and a persistent component in the conditional mean of the aggregate dividend growth rate. The diagonal elements of the transition probability matrix are 0.7 and 0.8, respectively, implying that the mean duration of the two regimes are about 3 and 5 years, respectively. The implied unconditional probabilities of being in the two regimes are 0.4 and 0.6, respectively. Thus, even though there is a very persistent, low frequency component in the consumption and dividend growth rates, the above feature of the data makes it difficult to capture it in single-regime models. This reconciles two strands of literature, one of which has argued that consumption growth is an i.i.d. process and the other has emphasized the presence of a small, persistent expected growth rate component.

The second row of Panel A reports the Dist-stat to test the null hypothesis that the model is correctly specified. The statistic takes the value 0.001 which has an asymptotic p-value of 1%. Thus, we fail to reject the time-series specification at the 1% level of significance.

The results in Panel B for the efficient weighting matrix are almost identical to those in Panel A. The persistence parameter of the conditional mean of consumption growth and the level of its volatility in the two regimes takes the same values as in Panel A: 0.92 and 1.2%, respectively, in regime 0 and 0.50 and 2.1%, respectively, in regime 1. The J-stat has an asymptotic chi-squared distribution with 2 degrees of freedom under the null that the model is correctly specified. The J-stat takes the value 8.53 and has an asymptotic p-value of 1.4%. Thus, we fail to reject the time-series specification at the 1% level of significance.

The small p-values of the Dist-stat and the J-stat in Panels A and B of Table 2 may well be due to the well known tendency of GMM estimators to over-reject in finite samples. We perform Monte Carlo simulations to obtain the finite-sample critical values of the Dist-stat and the J-stat for over-identifying restrictions. The details of the simulation design are in Appendix A.7. We calibrate the parameters of the time series processes to their SME point estimates in Table 2 and set the initial values of the state variables to their unconditional means, $x_0 = 0$ and $p_0 = \frac{1-\pi_1}{2-\pi_0-\pi_1}$. Given the initial value, p_0, we simulate a time-series of p_t of the same length as the historical sample. We use this simulated time-series to draw the state of the economy, i.e. the regime, at each time period, and, hence, simulate the time-series of the conditional mean of consumption growth, and the aggregate consumption and dividend growth rates. We then perform the SME estimation of the 11 time-series parameters using the 13 time-series moment restrictions used in Table 2 and obtain the Dist-stat, D^i, and the J-stat, J^i. This procedure is repeated 100 times. The 90%, 95%, and 99% finite-sample critical values of the Dist-stat and the J-stat are obtained from the percentiles of $\{D^i\}_{i=1}^{100}$ and $\{J^i\}_{i=1}^{100}$, respectively.

The 90%, 95%, and 99% critical values from the finite-sample distribution of the Dist-stat are 0.01, 0.02, and 0.05, respectively. Thus, the Dist-stat obtained in Table 2, Panel A using historical data has a finite-sample p-value far exceeding 10%. Inference using the asymptotic distribution of the Dist-stat in Table 2 implies that we fail to reject the time-series model at the 1% level of significance. Using the small sample distribution raises the p-value of the statistic to more than 10% thereby providing much stronger evidence in favour of the specification. Similar conclusions are reached for the efficient weighting matrix. The 90%, 95%, and 99% critical values from the finite-sample distribution of the J-stat are 30.23, 35.22, and 53.91, respectively. Thus, the J-stat obtained in Table 2, Panel B using historical data has a finite-sample p-value exceeding 10%. To summarize, we fail to reject the time-series specification of the model in equations (1), (3), and (6) using both the identity and efficient weighting matrices and both the asymptotic and finite-sample critical values of the test statistics for over-identifying restrictions.

5.2 Empirical Evidence on the Equity Premium

We next examine the ability of the model to jointly explain the pricing restrictions given by the Euler equations for a set of assets and the restrictions on the unconditional moments of consumption and aggregate dividend growth rates implied by the time-series specification of the model. We first consider the case when the asset menu consists of the market portfolio and the risk free rate. The lagged log price-dividend ratio of the market and the lagged log risk free rate are used as instruments. The Euler equations for the two assets along with the two chosen instruments gives 6 moment restrictions. To this set of pricing restrictions, we add moment restrictions implied by the time-series specification of the model. In particular, as in Section 5.1, we include moments corresponding to the unconditional means, variances, and first, second, and third-order autocovariances of consumption and dividend growth rates, the covariance between consumption and dividend growth rates, and the covariances between consumption growth and one and two lags of the dividend growth rate. This gives 13 moment restrictions corresponding to the assumed time-series processes. Thus, we have a total of 19 moment conditions. The total number of parameters to be estimated is 14, including 11 time-series parameters - μ, μ_d, ϕ, φ_d, ρ_0, ρ_1, φ_e, σ_0, σ_1, π_0, π_1 - and 3 preference parameters - δ, γ, ψ.

Note that the Euler equations deliver moment restrictions that have analytical representations in terms of observable variables - the aggregate consumption growth rate, the market-wide price-dividend ratio, the risk free rate, and the returns on the set of test assets - and the unknown time-series and preference parameters, but the time-series moment restrictions do not have analytic representations. In this case, we adopt a hybrid estimation methodology that combines features of the GMM and SME approaches. This requires a small adjustment to the asymptotic distributional results in Duffie and Singleton (1993). The details of the methodology and construction of standard errors of the parameter estimates and the test statistics for testing the over-identifying restrictions are described in Appendix A.6. We perform the estimation and tests using both the identity and the efficient weighting matrices.[8]

---

[8] The identity weighting matrix puts equal weight on all the moment restrictions and arguments advanced in favour of it are its superior finite-sample properties and that it forces the estimates to minimize the sum of squared pricing errors. However, note that it ignores the covariance structure of the moment restrictions often leading to an identification problem. This issue is particularly serious

Note that the moment conditions corresponding to the pricing restrictions are highly nonlinear in the parameters making optimization difficult. In order to get accurate estimates, we adopt the following algorithm. In Section 5.1, we estimate the 11 time-series parameters using 13 moment restrictions on the unconditional moments of aggregate consumption and dividend growth rates implied by the time-series specification of the model. This procedure gives initial consistent estimates of these parameters along with their standard errors. Next, we obtain final estimates of the time-series parameters and also estimate the preference parameters by performing a 14-dimensional grid search, over the 11 time series parameters and 3 preference parameters, using the full set of 19 pricing and time-series restrictions. The grids for the persistence parameter of the conditional mean of consumption and dividend growth rates in the two regimes are $0.90, 0.92, ..., 0.98$, and $0.5, 0.6, ..., 0.9$, respectively. The grids for the volatility of the conditional mean in the two regimes are $0.006, 0.009, ..., 0.015$, and $0.012, 0.015, ..., 0.024$, respectively. For the other time series parameters, the grid consists of evenly spaced points within two standard errors of the initial consistent point estimates. The grid for the risk aversion parameter is $2, 4, ..., 10$, that for the IES is $0.6, 0.9, 1.2, 1.5$, and that for the rate of time-preference is $0.90, 0.91, ..., 0.99$.

Note that our estimation methodology involves inversion of two non-linear equations to express the latent state variables, x_t and p_t, as functions of the observables, $z_{m,t}$ and $R_{f,t}$. This procedure yields quadratic equations for x_t and p_t, with coefficients that depend on $z_{m,t}$ and $R_{f,t}$, and the time-series and preference parameters. Solving the equations gives two pairs of solutions for x_t and p_t as functions of the observables, $z_{m,t}$ and $R_{f,t}$.

Table 3 reports results obtained using the bigger root of the quadratic equations.[9] Panel A and B report results for the identity and the efficient weighting matrices, respectively. The first row of Panel A reports the point estimates of the parameters along with the associated standard errors in parentheses. The persistence parameter of the conditional mean of consumption growth in the two regimes takes values 0.94 and 0.60, respectively. This suggests that in regime 0, the conditional mean of consumption and dividend growth has a half-life of just over 11 years, i.e. it has a frequency much longer than the average period of the business cycle. In regime 1, consumption and dividend dynamics are driven by a much higher frequency component that has a half-life of just over 1 year. The volatility of the conditional mean of consumption growth takes values 0.9% and 2.4%, respectively, in the two regimes. These findings are qualitatively similar to those in Table 2 in that they suggest the presence of two regimes, one in

---

in our setting where the moment restrictions include the pricing restrictions as well as time-series restrictions. Using the efficient weighting matrix avoids the identification issues by taking into account the covariance structure of the moment conditions and appropriately weighting them. Thus, although we report results for both choices of the weighting matrix, the results obtained using the efficient weighting matrix are more reliable in this setting and, hence, focussed upon.

[9] We report results for the bigger root as this choice minimizes the value of the criterion function.

which consumption and dividend growth rates are more persistent and less volatile and the other during which the growth rates are much less persistent and have higher volatility.

In Table 2, using only time series data on aggregate consumption and dividend growth, we earlier estimated the volatility of the conditional mean of consumption growth in the two regimes as 1.2% and 2.1%, respectively. In Table 3, using the pricing restrictions in addition to the time series data on aggregate consumption and dividend growth makes the difference between the two regimes more pronounced. The estimated volatility in the first regime is lowered to 0.9% and in the second regime is raised to 2.4%.

The third and fourth rows of Panel A report the average pricing errors for the market portfolio and the risk free rate and the associated standard errors.[10] The average pricing errors are small at 3.5% and -2.9%, respectively, and are not statistically significantly different from zero. The last row of Panel A gives the Dist-stat for testing over-identifying restrictions. The Dist-stat is 1.37 and has an asymptotic p-value smaller than 1%.

The results in Panel B for the efficient weighting matrix are largely similar to those in Panel A. The persistence parameter of the conditional mean of consumption growth in the two regimes takes values 0.96 and 0.70, respectively. This suggests that the conditional mean of consumption and dividend growth has half-lives of about 17 years and 2 years, respectively, in the two regimes. The volatility of the conditional mean takes values 0.9% and 1.5%, respectively, in the two regimes. The average pricing errors for the market portfolio is somewhat large at 8.1%. However, it is not statistically distinguishable from zero. The average pricing error for the risk free rate is miniscule at 0.1%. The last row of the panel gives the J-stat for testing over-identifying restrictions. The J-stat is 24.53 and has an asymptotic p-value smaller than 1%.

Note that the non-linearity of the Euler equations with respect to the parameters, the large number of parameters to be estimated (14), and the relatively small sample size (77) calls into question the accuracy of the asymptotic inference in Table 3. Moreover, as noted in Section 5.1, the finite-sample critical values of the Dist-stat and the J-stat are several times bigger than the corresponding asymptotic ones. Hence, rejection of the model in Panels A and B of Table 3 using the asymptotic distribu-

---

[10] The average pricing error for asset j is computed as

$$\left[\frac{1}{T}\sum_{t} exp\left(m_{t+1}\left(\widehat{\Theta}\right)+r_{j,t+1}\right)-1\right].$$

where $\widehat{\Theta}$ denotes the point estimates of the model parameters. The standard error of the average pricing error is computed as $se\left(exp\left(m_{t+1}\left(\widehat{\Theta}\right)+r_{j,t+1}\right)-1\right)/\sqrt{T}$, where se denotes standard error. Note that, under the model assumptions, $exp\left(m_{t+1}\left(\widehat{\Theta}\right)+r_{j,t+1}\right)-1$ is a martingale difference sequence and, hence, the above procedure gives valid standard errors for the average pricing error.

tions of the Dist-stat and the J-stat may be misleading. We perform Monte Carlo simulations to examine the finite-sample performance of the estimators and obtain the finite-sample critical values of the Dist-stat and the J-stat for overidentifying restrictions. The details of the simulation design are in Appendix A.7. We calibrate the time-series and preference parameters to their SME point estimates in Table 3 and set the initial conditions of the state variables to their unconditional means, $x_0 = 0$ and $p_0 = \frac{1-\pi_1}{2-\pi_0-\pi_1}$. Given the initial value, p_0, we simulate a time-series of p_t of the same length as the historical sample. We use this simulated time-series to draw the state of the economy, i.e. the regime, at each time period, and, hence, simulate the time-series of the conditional mean of consumption growth, and the aggregate consumption and dividend growth rates. We simulate the time-series of log returns on the market portfolio and the risk free rate, using the log-linearization in equation (11) and the model solution in equation (14), respectively. We then perform the SME estimation of the parameters with the 19 pricing and time-series restrictions used in Table 3 and obtain the Dist-stat, D^i, and the J-stat, J^i. This procedure is repeated 100 times. The 90%, 95%, and 99% finite-sample critical values of the Dist-stat and the J-stat are obtained from the percentiles of $\{D^i\}_{i=1}^{100}$ and $\{J^i\}_{i=1}^{100}$, respectively.

The 90%, 95%, and 99% critical values from the finite-sample distribution of the Dist-stat are 0.48, 0.59, and 0.95, respectively. Thus, the Dist-stat obtained in Table 3, Panel A using historical data has a finite-sample p-value smaller than 1%. Inference using the asymptotic distribution of the Dist-stat in Table 3 implies that we reject the model at the 1% level of significance. Using the small sample distribution for inference also leads to rejection of the model; however, the finite-sample critical values are an order of magnitude bigger than the corresponding asymptotic ones. The 90%, 95%, and 99% critical values from the finite-sample distribution of the J-stat are 33.29, 41.78, and 54.64, respectively. Thus, the J-stat obtained in Table 3, Panel B using historical data has a finite-sample p-value bigger than 10%. Hence, although we reject the hypothesis that the model specification jointly explains the pricing restrictions as well as the restrictions on the unconditional moments of aggregate consumption and dividend growth rates using asymptotic critical values, inference using the finite-sample critical values evades rejection of the model.

5.3 Empirical Evidence on the Cross-Section of Returns

We next explore the ability of the model to explain the cross-section of annual asset returns over the period 1930-2006. The asset menu consists of the market portfolio, the risk free rate, and portfolios of "Value", "Growth", "Small" capitalization, and "Large" capitalization stocks as detailed in Section 4. The Euler equations for these 6 assets along with the 13 time-series moment restrictions gives 19 moment restrictions in 14 parameters. The optimization algorithm used is similar to that in Section 5.2.

Table 4 reports the estimation and test results obtained using the bigger root of the quadratic equations. Panel A and B report results for the identity and the efficient weighting matrices, respectively. The first row of Panel A reports the point estimates of the parameters along with the associated standard errors in parentheses. The persistence parameter of the conditional mean of consumption growth in the two regimes takes values 0.90 and 0.80, respectively. This suggests that in regime 0, the conditional mean of consumption and dividend growth has a half-life of about 7 years, while in regime 1 it has a half-life of just over 3 years. The volatility of the conditional mean of consumption growth takes values 1.1% and 2.5%, respectively, in the two regimes. These findings are qualitatively similar to those in Table 2 in that they suggest the presence of two regimes, one in which consumption and dividend growth rates are more persistent and less volatile and the other during which the growth rates are much less persistent and have higher volatility. The estimated values of the risk aversion and the IES parameters are 10 and 0.6, respectively. These are the same as the estimates obtained in Table 3 for the 2-asset system.

Panel A also report the average pricing errors for the 6 assets and the associated standard errors. The average pricing errors vary from -7.5% for the risk free rate to 8.6% for the portfolio of "Small" capitalization stocks. Although the magnitudes of the pricing errors are large, none of them is statistically significantly different from zero. The last row of Panel A gives the Dist-stat for testing the over-identifying restrictions. The Dist-stat is 1.28 and has an asymptotic p-value smaller than 1%.

Panel B reports results for the efficient weighting matrix. The persistence parameter of the conditional mean of consumption growth in the two regimes takes values 0.90 and 0.60, respectively. This suggests that the conditional mean of consumption and dividend growth has half-lives of about 7 years and just over 2 years, respectively, in the two regimes. The volatility of the conditional mean takes values 1.1% and 2.5%, respectively, in the two regimes. The average pricing errors for the "Large" capitalization portfolio, "Growth" portfolio and the market are statistically and economically very small at 0.2%, -0.3%, and 1.2%, respectively. The average pricing errors for the "Small" capitalization portfolio, "Value" portfolio and the risk free rate are economically somewhat large at 10.7%, 8.9%, and -7.1%, respectively. However, all of them are statistically indistinguishable from zero. The last row of the panel gives the J-stat for testing over-identifying restrictions. The J-stat is 49.95 and has an asymptotic p-value smaller than 1%.[11]

As in Sections 5.1. and 5.2, we perform Monte Carlo simulations to examine the finite-sample performance of the estimators and obtain the finite-sample critical values of the Dist-stat and the J-stat for overidentifying restrictions. The details of the simulation design are in Appendix A.7. We calibrate the time-series and preference parameters to their SME point estimates in Table 4 and set the initial conditions of the

---

[11] As in Section 5.2, we report results for the bigger root as this choice minimizes the value of the criterion function.

state variables to their unconditional means, $x_0 = 0$ and $p_0 = \frac{1-\pi_1}{2-\pi_0-\pi_1}$. Given the initial value, p_0, we simulate a time-series of p_t of the same length as the historical sample. We use this simulated time-series to draw the state of the economy, i.e. the regime, at each time period, and, hence, simulate the time-series of the conditional mean of consumption growth, and the aggregate consumption and dividend growth rates. We simulate the time-series of log returns on the market portfolio and the risk free rate, using the log-linearization in equation (11) and the model solution in equation (14), respectively. We simulate the series for the log returns on the Small, Large, Growth, and Value portfolios, using similar log-linearizations as for the market portfolio. We then perform the SME estimation of the parameters with the 19 pricing and time-series restrictions used in Table 4 and obtain the Dist-stat, D^i, and the J-stat, J^i. This procedure is repeated 100 times. The 90%, 95%, and 99% finite-sample critical values of the Dist-stat and the J-stat are obtained from the percentiles of $\{D^i\}_{i=1}^{100}$ and $\{J^i\}_{i=1}^{100}$, respectively.

The 90%, 95%, and 99% critical values from the finite-sample distribution of the Dist-stat are 1.33, 1.70, and 2.05, respectively. Thus, the Dist-stat obtained in Table 4, Panel A using historical data has a finite-sample p-value exceeding 10%. Inference using the asymptotic distribution of the Dist-stat in Table 4 implies that we reject the model at the 1% level of significance. Using the small sample distribution for inference raises the p-value of the statistic to more than 10% thereby providing much stronger evidence in favour of the specification. Similar conclusions are reached for the efficient weighting matrix. The 90%, 95%, and 99% critical values from the finite-sample distribution of the J-stat are 33.44, 40.05, and 53.72, respectively. Thus, the J-stat obtained in Table 4, Panel B using historical data has a finite-sample p-value exceeding 1%. To summarize, although we reject the hypothesis that the model specification jointly explains the pricing restrictions as well as the restrictions on the unconditional momentsof aggregate consumption and dividend growth rates using asymptotic critical values for both the identity and efficient weighting matrices, using finite-sample critical values provides much stronger evidence in favour of the model.

5.4 Extracted State Variables: An Economic Interpretation of the Regimes

The regime shifts model presented in this paper involves two state variables - the conditional mean of consumption growth, x_t, and the posterior probability of being in the first regime, p_t. The representative consumer observes the conditional mean and forms the posterior probability at time t of being in regime $s_t = 0$ based on the history of consumption, dividends, the conditional mean of consumption growth, and any other information that he uses to form expectations. The econometrician's information set is generally a subset of the consumer's information set and, hence, the probability, p_t, and the conditional mean of consumption growth, x_t, are latent state variables.

A nice feature of the model is that the latent state variables become observable to the econometrician because, in the log-linearized version of the model, the aggregate log price-dividend ratio and the risk free rate are affine functions only of the two state variables and their product. The coefficients are known functions of the underlying time-series and preferences parameters. Hence, we can invert the system to extract the two state variables as known functions of the observable aggregate log price-dividend ratio and interest rate. This procedure of inverting two non-linear equations yields quadratic equations for x_t and p_t, with coefficients that depend on $z_{m,t}$ and $R_{f,t}$, and the time-series and preference parameters. Solving the equations gives two pairs of solutions for x_t and p_t as functions of the observables, $z_{m,t}$ and $R_{f,t}$. We use the bigger root of the quadratic equations to extract the latent state variables. This is because, as pointed out in Sections 5.2 and 5.3, this choice minimizes the value of the criterion function. The SME point estimates of the time-series and preference parameters are used in the extraction of the latent state variables.

Figures 1 and 2 report the estimated time-series of the posterior probability of being in the first regime, p_t, and the conditional mean of consumption growth, x_t, respectively, over the period 1931-2006, NBER recession periods (shaded areas) and the major stock market crashes identified by Mishkin and White (2002) plus the 2002 market crash (vertical dashed lines).[12] We classify a given year as a recession if a NBER recession was registered in at least one of the quarters.

Consider first Figure 1 that plots the extracted time-series of the posterior probability of being in the first regime. The first three years of the sample, 1931-1933, cover the period of the Great Depression. An NBER recession was registered over this period which was also characterized by major stock market crashes in each of the three years. The figure identifies this period with the first regime characterized by a short half-life and higher volatility of the conditional mean of consumption and dividend growth rates. Soon after the recession period, the economy moves to the secong regime characterized by a half-life of more than 3 years and lower volatility of the conditional mean of consumption and dividend growth rates. It moves back to the first regime during the 1937-1938 recession period and stays there during the stock market crash in 1940. The expansionary period of the early forties is characterized by a near one posterior probability of the economy being in the secong regime with a half-life of more than 3 years and lower volatility of the conditional mean of consumption and dividend growth rates. The period 1948-1962, which includes four NBER registered recessions, is identified as corresponding to the first regime. The economy gradually moves to the second regime during the expansionary phase of the mid and late sixties. The period

---

[12] Mishkin and White (2002) identify a stock market crash as a period in which either the Dow Jones Industrials, the S&P500 or the NASDAQ index drops by at least 20 percent in a time window of either one day, five days, one month, three months or one year.

1973-1992, which includes four NBER registered recessions and 4 major stock market crashed, is also identified as corresponding to the first regime. This period includes the largest one-day decline in stock market values in U.S. history –October 19th 1987, aka "Black Monday". Immediately after the 1992 recession, the economy moves to the second regime characterized by a half-life of more than 3 years and lower volatility of the conditional mean of consumption and dividend growth rates. It moves briefly to the first regime during the stock market crash in 1998 before reverting back to the second regime and staying there till the end of the sample period.

Thus, although the extraction of the latent state variables relies on the SME estimates of the model parameters and these are estimated with error, we identify two distinctly different patterns in the mean reversion of the conditional mean of the aggregate consumption and aggregate dividend growth rates. The first regime, characterized by a short (less than 3 years) half-life and higher volatility of the conditional mean of aggregate consumption and dividend growth rates, is more likely to correspond to NBER registered recession periods. The probability of being in a recession conditional on being in this regime is 0.41. The second regime, characterized by a half-life of more than 3 years and lower volatility of the conditional mean, seems to correspond to the expansionary phase of the business cycle. The probability of being in a recession conditional on being in this regime is only 0.23, almost half of that obtained for the first egime.

6 Predicting the Market Return, the Equity Premium, and the Cross-Section of Returns

6.1 In-Sample Forecasting Performance for the Market Return and the Equity Premium

We examine the ability of the regime shifts model to predict stock market return and the equity premium. Appendix A.4 shows that the continuously compounded return on the market portfolio implied by the model is a non-linear function of the two state variables, x and p,

$$E\left[r_{m,t+1}|\digamma(t)\right] = B_0 + B_1 x_t + B_2 p_t + B_3 p_t x_t + B_4 p_t^2 x_t. \tag{19}$$

Hence, the realizations of the two state variables at period t, x_t and p_t, should predict the market return at time $t+1$. Using the time-series of the extracted state variables, we perform an ordinary least squares forecasting regression of the form suggested by equation (19),

$$r_{m,t+1} = \alpha_0 + \alpha_1 x_t + \alpha_2 p_t + \alpha_3 p_t x_t + \alpha_4 p_t^2 x_t + \epsilon_{t+1}^1. \tag{20}$$

The *adjusted-*R^2 from the regression provides a measure of the forecasting performance of the model.

We use a similar procedure to examine the predictive performance of the model for the equity premium. Appendix A.4 shows that the equilibrium equity premium implied by the model is a non-linear function of the two state variables, x and p,

$$E\left[\left(r_{m,t+1} - \frac{1}{R_{f,t}}\right)|\mathcal{F}(t)\right] = E_0 + E_1 x_t + E_2 p_t + E_3 p_t x_t + E_4 p_t^2 x_t. \tag{21}$$

Hence, using the time-series of the extracted state variables, we perform an ordinary least squares forecasting regression of the form suggested by equation (21),

$$r_{m,t+1} - \frac{1}{R_{f,t}} = \alpha_0 + \alpha_1 x_t + \alpha_2 p_t + \alpha_3 p_t x_t + \alpha_4 p_t^2 x_t + \epsilon_{t+1}^2. \tag{22}$$

The results of the forecasting regression for the full sample period 1930-2006 are presented in Table 5. Panel A reports results for the market portfolio while Panel B does the same for the equity premium. The first row of Panel A reports results of the regression in equation (20). Note that the regression coefficients of the two state variables are statistically significant. The *adjusted-*R^2 of the regression is 7.3%.

Note that the two state variables, x and p, are non-linear functions of the aggregate log price-dividend ratio and the gross interest rate (see equations (13) and (14)). In other words, the model predicts a complicated non-linear relationship between the equilibrium market return and the aggregate log price-dividend ratio and gross interest rate. Rows 2 and 3 of Panel A show the loss in forecasting power that results from ignoring this non-linearity and performing a linear regression of continuously compounded market return on lagged log price-dividend ratio and gross risk free rate. Linear predictive regressions with these choices of predictor variables are among the most popular in the predictability literature to predict stock market returns and the equity premium. Row 2 reports results from a linear regression of market returns on the lagged aggregate log price-dividend ratio. Both the constant and the coefficient on the price-dividend ratio are statistically significant. Moreover, the latter has the expected negative sign. However, the *adjusted-*R^2 is only 3.5% - less than half of that obtained from the model-implied regression in Row 1. Row 3 reports results from a linear regression of market return on the lagged aggregate log price-dividend ratio and the gross risk free rate. Note that only the coefficient on the price-dividend ratio is statistically significant while the constant and the coefficient on the risk free rate are not significantly different from zero. The *adjusted-*R^2 rises marginally to 3.9% and is still only about half of that obtained from the model-implied regression.

The results in Panel B for the equity premium are largely similar to those in Panel A. The first row reports results of the model-implied regression in equation (22). Note that the constant and the regression coefficient of the state variable denoting the economic regime are statistically significant. The *adjusted-*R^2 of the regression is 10.3%. Row 2 reports results from a linear regression of the realized equity premium (defined by the left-hand-side of equation (22)) on the lagged aggregate log price-dividend ratio. The *adjusted-*R^2 is only 1.0% - an order of magnitude smaller than that obtained from the model-implied regression in Row 1. Row 3 reports results from a linear regression of the realized equity premium on the lagged aggregate log price-dividend ratio and the gross risk free rate. Note that all the regression coefficients are statistically significantly different from zero. The *adjusted-*R^2 rises to 7.3% but is still much smaller than that obtained from the model-implied regression.

The forecasting performance of the model improves further when attention is restricted to the post war subperiod 1947-2006. The results of the forecasting regressions for the post war subperiod are presented in Table 6. Panel A reports results for the market portfolio while Panel B does the same for the equity premium. The first row of Panel A reports results of the model-implied regression for the stock market return in equation (20). The *adjusted-*R^2 of the regression rises to 15.0% - a two-fold increase compared to a value of 7.3% obtained for the full sample period. Row 2 of Panel A reports results from a linear regression of continuously compounded market return on the lagged log price-dividend ratio over the post war sample period. The *adjusted-*R^2 is only 5.5% - more than five times smaller than that obtained from the model-implied regression in Row 1. Row 3 reports results from a linear regression of market return on the lagged aggregate log price-dividend ratio and the gross risk free rate. The *adjusted-*R^2 rises to 8.0% but is still more than three times smaller than that obtained from the model-implied regression.

The results in Panel B for the equity premium are largely similar to those in Panel A. The first row reports results of the model-implied regression in equation (22). The *adjusted-*R^2 of the regression is 19.6%, almost double of that obtained in Table 5 for the full sample period. Row 2 reports results from a linear regression of the realized equity premium on the lagged aggregate log price-dividend ratio. The *adjusted-*R^2 is only 4.1% - almost five times smaller than that obtained from the model-implied regression in Row 1. Row 3 reports results from a linear regression of the realized equity premium on the lagged aggregate log price-dividend ratio and the gross risk free rate. The *adjusted-*R^2 rises to 12.0% but is still much smaller than that obtained from the model-implied regression.

Overall, the regime shifts model performs considerably better in-sample at predicting the returns on the market portfolio and the equity premium than linear forecasting regressions of returns on lagged price-dividend ratio and risk free rate that are widely employed in the predictability literature. This result holds for the full sample period 1930-2006 as well as the post war subperiod 1947-2006, thereby lending additional support to the hypothesis of regime shifts highlighted in the paper.

6.2 Out-of-Sample Forecasting Performance for the Market Return and the Equity Premium

We also examine the ability of the regime shifts model to predict stock market return and the equity premium out-of-sample over the forty year period 1967-2006. At each year, all available data prior to that date is used to estimate equations (20) and (22). The estimates of the regression coefficients are used to forecast the market return and the equity premium in the next year. The choice of the sample period 1967-2006 ensures that a sufficient number of observations are used in the initial estimation to obatin reliable estimates. The *out-of-sample-*R^2 provides a measure of the out-of-sample forecasting performance of the model. This is computed as

$$R^2 = 1 - \frac{var(ep_t^{data} - ep_t^{pred})}{var(ep_t^{data})} \tag{23}$$

for the equity premium and as

$$R^2 = 1 - \frac{var(mkt_t^{data} - mkt_t^{pred})}{var(mkt_t^{data})} \tag{24}$$

for the stock market return. In equation (23), ep_t^{data} denotes the realized equity premium in year t and ep_t^{pred} denotes the premium predicted by the model. Similarly, in equation (24), mkt_t^{data} denotes the realized stock market return in year t and mkt_t^{pred} denotes the return predicted by the model.

To facilitate comparison, we also report the out-of-sample performance of linear predictive regressions of market returns and the equity premium on the lagged log market-wide price-dividend ratio and the risk free rate.

The estimation results are presented in Table 7. Panel A reports results for the market portfolio while Panel B does the same for the equity premium. The first row of Panel A reports the out-of-sample results for the forecasting regression in equation (20). The *out-of-sample-*R^2 is 5.8%. Row 2 reports results from a linear regression of market returns on the lagged aggregate log price-dividend ratio. In this case, the *out-of-sample-*R^2 is only 0.5% - more than 10 times smaller than that obtained from the model-implied regression in Row 1. Row 3 reports results from a linear regression of market return on the lagged aggregate log price-dividend ratio and the gross risk free rate. The *out-of-sample-*R^2 rises marginally to 1.4% and is still almost five times smaller than that obtained from the model-implied regression.

The results in Panel B for the equity premium lead to similar conclusions. The first row reports results of the model-implied regression in equation (22). The *out-of-sample-*R^2 of the predictive regression is economically substantial 10.9%. Row 2 reports results from a linear regression of the realized equity premium on the lagged aggregate log price-dividend ratio. The *out-of-sample-*R^2 is only 2.2% - almost five times smaller than that obtained from the model-implied regression in Row 1. Row 3 reports results from a linear regression of the realized equity premium on the lagged aggregate log price-dividend ratio and the gross risk free rate. The *out-of-sample-*R^2 rises to 6.9% but is still much smaller than that obtained from the model-implied regression.

6.3 In-Sample Forecasting Performance for the Cross-Section of Returns

We next examine the ability of the regime shifts model to predict the cross-section of returns of the size and book-to-market-equity-sorted portfolios. Forecastability of the cross-section of returns is important for at least two reasons. First, the historically observed magnitudes of the size and value premia, 4.5% and 4.1%, respectively, are almost as large as the equity premium, 5.8%, and, hence, they too have dramatic long term investment implications. Hence, successful identification of predictor variables has important implications for market timing. Second, asset pricing models typically attempt to explain, apart from the equity premium, other asset pricing anomalies like the size and value premia. Therefore, the state variables in these models should not only successfully forecast the aggregate market return, but also the returns on the size and the book-to-market-equity-sorted portfolios. This provides an alternative channel to examine the empirical plausibility of these models.

We examine the predictive power of the two state variables in our regime shifts model - the conditional mean of the aggregate consumption and dividend growth rates, x_t, and the consumer's posterior probability of being in the first regime at date t, p_t - for the "Small", "Large", "Growth", and "Value" portfolio returns by performing in-sample linear predictive regressions of the form

$$r_{i,t+1} = \alpha_i + \alpha_i x_t + \alpha_i p_t + \alpha_i p_t x_t + \alpha_i p_t^2 x_t + \epsilon_{t+1}^i, \quad i = s, l, g, v. \tag{25}$$

The *adjusted-*R^2 from the model implied predictive regressions in equation (25) provides a measure of the ability of the model to explain the cross-section of returns. We use the market-wide log price-dividend ratio and the gross risk free rate in the extraction of the state variables x_t and p_t as in the preceding two subsections. To facilitate comparison, we also report the in-sample performance of linear predictive regressions of the portfolio returns on the lagged log market-wide price-dividend ratio and the risk free rate.

The results of the forecasting regression for the full sample period 1930-2006 are presented in Table 8 for the size-sorted portfolios. Panel A reports results for the "Small" portfolio while Panel B does the same for the "Large" portfolio. The first row of Panel A reports results of the regression in equation (25). Note that the regression coefficients of the two state variables are statistically significant. The *adjusted-*R^2 of the regression is 4.9%. Row 2 reports results from a linear regression of the returns on the "Small" portfolio on the lagged aggregate log price-dividend ratio. Both the constant and the coefficient on the price-dividend ratio are statistically significant. The latter has the expected negative sign. However, the *adjusted-*R^2 is only 2.5% - half of that obtained from the model-implied regression in Row 1. Row 3 reports results from a linear regression of the return on the lagged aggregate log price-dividend ratio and the gross risk free rate. Note that only the coefficient on the price-dividend ratio is statistically significant while that on the risk free rate is not significantly different from zero. This is also reflected in the *adjusted-*R^2 the actually falls to 1.2% on the inclusion of the risk free rate.

The results in Panel B for the "Large" portfolio also show the superior forecasting performance of the regime shifts model. The first row reports results of the model-implied regression in equation (25). Note that the regression coefficients of the two state variables are statistically significant. The *adjusted-*R^2 of the regression is 7.8%. Row 2 reports results from a linear regression of the realized returns on the portfolio on the lagged aggregate log price-dividend ratio. The *adjusted-*R^2 is only 2.9% - almost three times smaller than that obtained from the model-implied regression in Row 1. Row 3 reports results from a linear regression of the realized portfolio returns on the lagged aggregate log price-dividend ratio and the gross risk free rate. The *adjusted-*R^2 rises to 3.8% but is still less than half of that obtained from the model-implied regression.

Table 9 presents the results of the forecasting regression for the full sample period 1930-2006 for the book–to-market-equity-sorted portfolios. Panel A reports results for the "Growth" portfolio while Panel B does the same for the "Value" portfolio. The first row of Panel A reports results of the regression in equation (25). The *adjusted-*R^2 of the regression is economically large at 10.5%. Row 2 shows that a linear regression of the returns on the "Growth" portfolio on the lagged aggregate log price-dividend ratio leads to an *adjusted-*R^2 is only 2.3% while Row 3 shows that inclusion of the gross risk free rate as an additional predictor variable actually lowers the *adjusted-*R^2 to 2.0%.

The results in Panel B for the "Value" portfolio show the regime shifts model does not do very well for this portfolio. The model-implied regression in Row 1 has an *adjusted-*R^2 of only 0.5%. A linear regression of the realized returns on the portfolio on the lagged aggregate log price-dividend ratio gives an *adjusted-*R^2 of 2.7%.

The forecasting performance of the model improves further when attention is restricted to the post war subperiod 1947-2006. The results of the forecasting regressions for the post war subperiod are presented in Table 10 and 11 for the size and the book-to-market-equity-sorted portfolios, respectively. The *adjusted-*R^2 for the "Small", "Large", "Growth", and "Value" portfolio returns from the model-implied predictive regressions are 4.6%, 15.4%, 18.8%, and 3.2%, respectively. These values are much larger than those obtained from linear predictive regressions of realized returns on the lagged log market price-dividend ratio and the risk free rate (the *adjusted*-R^2 from the latter regressions are -1.8%, 9.6%, 5.1%, and 2.6%, respectively).

7 Conclusion and Extensions

We present an asset pricing model with two regimes where the conditional mean of the aggregate consumption growth rate reverts to its unconditional mean with speed and variance which differ across regimes. The representative consumer, endowed with recursive preferences, observes the conditional mean, but not the regime. At each period, he forms a posterior probability of being in the first regime using his current information set. The econometrician's information set is a subset of that of the consumer. Hence, the two state variables - the conditional mean of the aggregate consumption growth rate and the consumer's posterior probability of being in the first regime - are latent to the econometrician. We show that, in equilibrium, the market-wide log price-dividend ratio and the risk free rate, and, hence, the pricing kernel, are affine functions of the two latent state variables and their product.

We examine the empirical plausibility of the model using the SMM approach of Duffie and Singleton (1993). The stochastic discount factor in the economy is a function of the two state variables and their cross product. Since the market-wide log price-dividend ratio and the risk free rate are affine functions of the state variables and their product, these two equations may be inverted to extract the two latent state variables as known functions of the observable log price-dividend ratio and the risk free rate. Hence, we rewrite the pricing kernel as a function of the observable log price-dividend ratio and the risk free rate. We substitute this expression for the pricing kernel into the set of Euler equations to obtain a set of moment restrictions that are expressed entirely in terms of observables. Also, the time-series specification of the model imposes constraints on the unconditional moments of the aggregate consumption and dividend growth rates. These are additional constraints, in addition to the pricing restrictions.

We first examine the empirical plausibility of the time-series specification of the model. Using aggregate consumption and dividend data alone, and the constraints imposed on the unconditional moments of the aggregate consumption and dividend growth rates by the time-series specification of the model, we identify two distinctly different patterns in the mean reversion of the conditional mean of the aggregate consumption and aggregate dividend growth rates. In the first regime, the half-life of the conditional mean is shorter than three years, while in the second regime, it is longer than three years. The volatility of the conditional mean in the first regime is about double of that in the second regime. Moreover, the over-identifying restrictions test fails to reject the time-series specification of the model.

Next we test the ability of the model to jointly explain the pricing restrictions, given by the Euler equations for a set of assets, and the restrictions on the unconditional moments of aggregate consumption and dividend growth rates implied by the time-series specification of the model. We find that the two-regime model produces significantly lower pricing errors in the market portfolio, the risk free rate, and the cross-section of asset returns over the entire available sample period 1930-2006 than the single-regime model does and that the over-identifying restrictions are not rejected.

The time-series of the extracted state variables reveal that the first regime, characterized by a short (less than 3 years) half-life and higher volatility of the conditional mean of aggregate consumption and dividend growth rates, is more likely to correspond to NBER registered recession periods. The second regime, characterized by a half-life of more than 3 years and lower volatility of the conditional mean, is more likely to correspond to the expansionary phase of the business cycle.

Finally, we examine the ability of the regime shifts model to predict stock market returns, the equity premium, and the cross-section of returns using as predictive variables the market-wide price-dividend ratio and the risk free rate. The model has superior predictive performance for the market returns, the equity premium and the cross-section of returns compared to linear predictive regressions of returns on the lagged aggregate price-dividend ratio and the risk free rate, both in-sample and out-of-sample. These findings lend additional support to the regime shifts model highlighted in the paper. In future research, we plan to model the dividend growth processes of different classes of assets and explore further issues of predictability.

References

[1] Alvarez, F., and Jermann, U., (2005), "Using Asset Prices to Measure the Persistence of the Marginal Utility of Wealth," *Econometrica*, 73(6), 1977-2016.

[2] Bansal, R., Dittmar, R. F., and Lundblad, C., (2005), "Consumption, Dividends, and the Cross-Section of Equity Returns," *Journal of Finance* 60, 1639-1672.

[3] Bansal, R., Gallant, R., and Tauchen, G., (2007), "Rational Pessimism, Rational Exhuberance, and and Asset Pricing Models," *Review of Economic Studies*, 74, 1005-1033.

[4] Bansal, R., Kiku, D., and Yaron, A., (2007), "Risks For the Long Run: Estimation and Inference," Working Paper.

[5] Bansal, R., and Yaron, A., (2004), "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, 1481-1509.

[6] Bekaert, G., Engstrom, E., and Xing, Y., (2005), "Risk, Uncertainty, and Asset Prices," Working Paper.

[7] Blanchard, O. J., and Simon, J., (2001), "The Long and Large Decline in U.S. Output Volatility," *Brookings Papers on Economic Activity*, 1, 135–174.

[8] Campbell, J. Y., (2000), "Asset Pricing at the Millennium," *Journal of Finance*, 55(4), 1515-1567.

[9] Campbell, J. Y., (2003), "Consumption-Based Asset Pricing," in *Handbook of the Economics of Finance, Vol. I-B*, ed. by G. M. Constantinides, M. Harris, and R. Stulz, pp. 803-887. North Holland, Amsterdam.

[10] Campbell, J. Y., Lo, A. W., and MacKinlay, C., (1997), "*The Econometrics of Financial Markets*," Princeton University Press, Princeton, NJ.

[11] Campbell, J. Y., and Shiller, R.J., (1988), "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies*, 1, 195-227.

[12] Cochrane, J. H., (2005), "*Asset Pricing*," Revised Edition. Princeton University Press, Princeton, NJ.

[13] Cochrane, J. H., and Hansen, L.P., (1992), "Asset Pricing Explorations for Macroeconomics," *Macroeconomics Annual.*

[14] Constantinides, G. M., (2002), "Rational Asset Prices," *Journal of Finance*, 57(4), 1567-1591.

[15] Constantinides, G. M., and Ghosh, A., (2008), "Asset Pricing Tests with Long Run Risks in Consumption Growth," *Working paper.*

[16] Duffie, D., and Singleton, K.J., (1993), "Simulated Moments Estimation of Markov Models of Asset Prices," *Econometrica,* 61(4), 929-952.

[17] Epstein, L. R., and Zin, S. E., (1989), "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica,* 57, 937-969.

[18] Fama, E. F., and French, K. R., (1993), "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics*, 33, 3-56.

[19] Hansen, L. P., (1982), "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029-1054.

[20] Hansen, L.P., Heaton, J., and Li, N., (2008), "Consumption Strikes Back," *Journal of Political Economy*, 116(2), 260-302.

[21] Hansen, L. P., and Scheinkman, J., (2007), "Long-Term Risk : An Operator Approach," Working Paper.

[22] Jagannathan, R., and Wang, Z., (1996), "The Conditional CAPM and the Cross Section of Expected Returns," *Journal of Finance*, 51, 3-53.

[23] Kiku, D., (2006), "Long Run Risks and the Value Premium Puzzle," Working Paper, Duke University.

[24] Kim, C. J., and Nelson, C. R., (1999), "Has the Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle," *The Review of Economics and Statistics*, 81, 608–616.

[25] Kocherlakota, N., (1996), "The Equity Premium: It's Still A Puzzle," *Journal of Economic Literature*, 34(1), 42-71.

[26] Koijen, R. S. J., and Van Binsbergen, J. H., (2008), "Predictive Regressions: A Present-Value Approach," *Working paper.*

[27] Kreps, D. M., and Porteus, E. L., (1978), "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica*, 46(1), 185-200.

[28] Lettau, M., and Ludvigson, S., (2008), "Euler Equation Errors," Working Paper, New York University.

[29] Lettau, M., Ludvigson, S., and Wachter, J., (2008), "The Declining Equity Premium: What Role Does Macroeconomic Risk Play?," *Review of Financial Studies*, 21(4), 1653-1687.

[30] Lettau, M., and Van Nieuwerburgh, S., (2008), "Reconciling the Return Predictability Evidence," *Review of Financial Studies*, 21(4), 1607-1652.

[31] Malloy, C., Moskowitz, T., and Vissing-Jorgensen, A., (2006), "Long-Run Stockholder Consumption Risk and Asset Returns," Working Paper, University of Chicago.

[32] McConnell, M. M., and Perez-Quiros, G., (2000), "Output Fluctuations in the United States: What Has Changed Since the early 1980s?," *American Economic Review*, 90(5), 1464–1476.

[33] Mehra, R., (2008), "*Handbook of the Equity Risk Premium*," ed., Elsevier, Amsterdam.

[34] Mehra, R., and Prescott, E. C., (2003), "The Equity Premium Puzzle in Retrospect," in *Handbook of the Economics of Finance, Vol. I-B*, ed. by G. M. Constantinides, M. Harris, and R. Stulz, pp. 803-887. North Holland, Amsterdam.

[35] Mishkin, F. S., and White, E. N., (2002), "U.S. Stock Market Crashes and Their Aftermath: Implications for Monetary Policy," *NBER Working Paper 8992*, National Bureau of Economic Research, Inc.

[36] Parker, J. A., and Julliard, C., (2005), "Consumption Risk and the Cross Section of Expected Returns," *Journal of Political Economy*, 113(1), 185-222.

[37] Pastor, L., and Stambaugh, R. F., (2001), "The Equity Premium and Structural Breaks," *Journal of Finance*, 56(4), 1207-1239.

[38] Stock, J. H., and Watson, M. W., (2002), "Has the Business Cycle Change and Why?," in *NBER Macroeconomics Annual*: 2002, ed. by M. Gertler, and K. Rogoff. MIT Press, Cambridge, MA.

[39] Weil, P., (1989), "The Equity Premium Puzzle and the Risk-Free Rate Puzzle," *Journal of Monetary Economics*, 24(3), 401-421.

A Appendix

Here, we derive the pricing implications of the regime shifts model.

A.1 Consumption Claim

We rely on the log-linear approximation for the continuous return on the consumption claim, $r_{c,t+1}$, as in Campbell and Shiller (1988),

$$r_{c,t+1} = \kappa_0 + \kappa_1 z_{t+1} - z_t + \Delta c_{t+1},$$

where z_t is the log price-consumption ratio. Note that the current model specification involves two latent state variables, x_t and p_t. We conjecture that the log price-consumption ratio at date t takes the form,

$$z_t = p_t \left[A_0(0) + A_1(0)x_t\right] + (1 - p_t)\left[A_0(1) + A_1(1)x_t\right].$$

The Euler equation for the consumption claim is,

$$E\left[\exp\left(m_{t+1} + r_{c,t+1}\right) | \mathcal{F}(t)\right] = 1, \tag{26}$$

$$m_{t+1} = \theta \log \delta - \frac{\theta}{\psi}\Delta c_{t+1} + (\theta - 1) r_{c,t+1}.$$

Substituting the expression for m_{t+1} from (9) into (26), we have,

$$E\left[\exp\left(\theta \log \delta - \frac{\theta}{\psi}\Delta c_{t+1} + \theta r_{c,t+1}\right) | \mathcal{F}(t)\right] = 1. \tag{27}$$

Using the law of iterated expectations, equation (27) implies,

$$E\left[E\left\{\exp\left(\theta \log \delta - \frac{\theta}{\psi}\Delta c_{t+1} + \theta r_{c,t+1}\right) | \mathcal{F}(t), s_{t+1}\right\} | \mathcal{F}(t)\right] = 1. \tag{28}$$

Using the conditional distribution of s_{t+1}, equation (28) may be written as,

$$\sum_{i=0}^{1} P\left(s_{t+1} = i | \mathcal{F}(t)\right) E\left\{\exp\left(\theta \log \delta - \frac{\theta}{\psi}\Delta c_{t+1} + \theta r_{c,t+1}\right) | \mathcal{F}(t), s_{t+1} = i\right\} = 1 \tag{29}$$

Now, consider the term $E\left\{\exp\left(\theta \log \delta - \frac{\theta}{\psi}\Delta c_{t+1} + \theta r_{c,t+1}\right) | \mathcal{F}(t), s_{t+1} = 0\right\}$. Using (1), (4), (10) and (12),

$$
\begin{aligned}
& E\left\{\exp\left(\theta\log\delta - \frac{\theta}{\psi}\Delta c_{t+1} + \theta r_{c,t+1}\right)|\mathcal{F}(t), s_{t+1} = 0\right\} \\
= \; & E[\exp\{\theta\log\delta - \frac{\theta}{\psi}\mu - \frac{\theta}{\psi}x_t - \frac{\theta}{\psi}\sigma_0\eta_{t+1} + \theta\kappa_0 + \theta\kappa_1 A_0(0) + \theta\kappa_1 A_1(0)\rho_0 x_t + \theta\kappa_1 A_1(0)\varphi_e\sigma_0 e_{t+1} \\
& -\theta p_t\left[A_0(0) + A_1(0)x_t\right] - \theta(1-p_t)\left[A_0(1) + A_1(1)x_t\right] + \theta\mu + \theta x_t + \theta\sigma_0\eta_{t+1}\}|p_t, x_t] \\
= \; & \exp\{\theta\log\delta + \left(-\frac{\theta}{\psi} + \theta\right)\mu + \theta\kappa_0 + \theta\kappa_1 A_0(0) - \theta A_0(1) \\
& + \left(-\frac{\theta}{\psi} + \theta\kappa_1 A_1(0)\rho_0 - \theta A_1(1) + \theta\right)x_t + \left(-\theta A_0(0) + \theta A_0(1)\right)p_t + \left(-\theta A_1(0) + \theta A_1(1)\right)p_t x_t \\
& + \frac{1}{2}\left[\left(-\frac{\theta}{\psi} + \theta\right)^2\sigma_0^2 + \left(\theta\kappa_1 A_1(0)\varphi_e\right)^2\sigma_0^2\right]\}.
\end{aligned}
$$

Similar calculations yield the expression for $E\left\{\exp\left(\theta\log\delta - \frac{\theta}{\psi}\Delta c_{t+1} + \theta r_{c,t+1}\right)|\mathcal{F}(t), s_{t+1} = 1\right\}$. Substituting these expressions into (29) and noting that $P\left(s_{t+1} = 0|\mathcal{F}(t)\right) = \left[\pi_0 p_t + \left(1-\pi_1\right)\left(1-p_t\right)\right]$, we have,

$$
\begin{aligned}
1 \; = \; & \left[\pi_0 p_t + \left(1-\pi_1\right)\left(1-p_t\right)\right] \times (\exp\{\theta\log\delta + \left(-\frac{\theta}{\psi} + \theta\right)\mu + \theta\kappa_0 + \theta\kappa_1 A_0(0) - \theta A_0(1) \\
& + \left(-\frac{\theta}{\psi} + \theta\kappa_1 A_1(0)\rho_0 - \theta A_1(1) + \theta\right)x_t + \left(-\theta A_0(0) + \theta A_0(1)\right)p_t + \left(-\theta A_1(0) + \theta A_1(1)\right)p_t x_t \\
& + \frac{1}{2}\left[\left(-\frac{\theta}{\psi} + \theta\right)^2\sigma_0^2 + \left(\theta\kappa_1 A_1(0)\varphi_e\right)^2\sigma_0^2\right]\}) \\
& + \left[1 - \pi_0 p_t - \left(1-\pi_1\right)\left(1-p_t\right)\right] \times (\exp\{\theta\log\delta + \left(-\frac{\theta}{\psi} + \theta\right)\mu + \theta\kappa_0 + \theta\kappa_1 A_0(1) - \theta A_0(1) \\
& + \left(-\frac{\theta}{\psi} + \theta\kappa_1 A_1(1)\rho_1 - \theta A_1(1) + \theta\right)x_t + \left(-\theta A_0(0) + \theta A_0(1)\right)p_t + \left(-\theta A_1(0) + \theta A_1(1)\right)p_t x_t \\
& + \frac{1}{2}\left[\left(-\frac{\theta}{\psi} + \theta\right)^2\sigma_1^2 + \left(\theta\kappa_1 A_1(1)\varphi_e\right)^2\sigma_1^2\right]\})
\end{aligned}
$$

Using the approximation $e^y - 1 \approx y$, we have,

$$\begin{aligned}
0 \;=\;& [\pi_0 p_t + (1-\pi_1)(1-p_t)] \times (\{\theta\log\delta + \left(-\frac{\theta}{\psi}+\theta\right)\mu + \theta\kappa_0 + \theta\kappa_1 A_0(0) - \theta A_0(1) \\
&+ \left(-\frac{\theta}{\psi} + \theta\kappa_1 A_1(0)\rho_0 - \theta A_1(1) + \theta\right) x_t + \left(-\theta A_0(0) + \theta A_0(1)\right) p_t + \left(-\theta A_1(0) + \theta A_1(1)\right) p_t x_t \\
&+\frac{1}{2}\left[\left(-\frac{\theta}{\psi}+\theta\right)^2 \sigma_0^2 + \left(\theta\kappa_1 A_1(0)\varphi_e\right)^2 \sigma_0^2\right]\}) \\
&+ [1-\pi_0 p_t - (1-\pi_1)(1-p_t)] \times (\{\theta\log\delta + \left(-\frac{\theta}{\psi}+\theta\right)\mu + \theta\kappa_0 + \theta\kappa_1 A_0(1) - \theta A_0(1) \\
&+ \left(-\frac{\theta}{\psi} + \theta\kappa_1 A_1(1)\rho_1 - \theta A_1(1) + \theta\right) x_t + \left(-\theta A_0(0) + \theta A_0(1)\right) p_t + \left(-\theta A_1(0) + \theta A_1(1)\right) p_t x_t \\
&+\frac{1}{2}\left[\left(-\frac{\theta}{\psi}+\theta\right)^2 \sigma_1^2 + \left(\theta\kappa_1 A_1(1)\varphi_e\right)^2 \sigma_1^2\right]\})
\end{aligned}$$

$$\begin{aligned}
=\;& \left[-\theta A_0(0) + \theta A_0(1)\right] p_t + \left[-\theta A_1(0) + \theta A_1(1)\right] p_t x_t + \left[\theta\log\delta + \left(-\frac{\theta}{\psi}+\theta\right)\mu + \theta\kappa_0 - \theta A_0(1)\right] \\
&+ [(1-\pi_1) + (\pi_0+\pi_1-1)p_t]\left\{\theta\kappa_1 A_0(0) + \frac{1}{2}\left[\left(-\frac{\theta}{\psi}+\theta\right)^2 \sigma_0^2 + \left(\theta\kappa_1 A_1(0)\varphi_e\right)^2 \sigma_0^2\right]\right\} \\
&+ [(1-\pi_1) + (\pi_0+\pi_1-1)p_t]\left(-\frac{\theta}{\psi} + \theta\kappa_1 A_1(0)\rho_0 - \theta A_1(1) + \theta\right) x_t \\
&+ [\pi_1 - (\pi_0+\pi_1-1)p_t]\left\{\theta\kappa_1 A_0(1) + \frac{1}{2}\left[\left(-\frac{\theta}{\psi}+\theta\right)^2 \sigma_1^2 + \left(\theta\kappa_1 A_1(1)\varphi_e\right)^2 \sigma_1^2\right]\right\} \\
&+ [\pi_1 - (\pi_0+\pi_1-1)p_t]\left(-\frac{\theta}{\psi} + \theta\kappa_1 A_1(1)\rho_1 - \theta A_1(1) + \theta\right) x_t
\end{aligned} \tag{30}$$

Now the Euler equation (30) must hold for all values of the state variables. Setting the coefficient of $p_t x_t$ to zero, we have,

$$\left[-\theta A_1(0) + \theta A_1(1)\right] + (\pi_0+\pi_1-1)\left\{\theta\kappa_1 A_1(0)\rho_0 - \theta\kappa_1 A_1(1)\rho_1\right\} = 0,$$

which implies,

$$A_1(0) = \frac{1-(\pi_0+\pi_1-1)\kappa_1\rho_1}{1-(\pi_0+\pi_1-1)\kappa_1\rho_0} A_1(1) \tag{31}$$

Setting the coefficient of x_t to zero gives,

$$(1-\pi_1)\left(-\frac{\theta}{\psi}+\theta\kappa_1 A_1(0)\rho_0-\theta A_1(1)+\theta\right)+\pi_1\left(-\frac{\theta}{\psi}+\theta\kappa_1 A_1(1)\rho_1-\theta A_1(1)+\theta\right)=0,$$

implying,

$$A_1(1)=\frac{1-\frac{1}{\psi}}{1-\pi_1\kappa_1\rho_1}+\frac{(1-\pi_1)\kappa_1\rho_0}{1-\pi_1\kappa_1\rho_1}A_1(0). \tag{32}$$

Solving equations (31) and (32) for $A_1(0)$ and $A_1(1)$ gives,

$$A_1(0)=\frac{\frac{1-\frac{1}{\psi}}{(1-\pi_1)\kappa_1\rho_0}}{\frac{1-(\pi_0+\pi_1-1)\kappa_1\rho_0}{1-(\pi_0+\pi_1-1)\kappa_1\rho_1}\frac{1-\pi_1\kappa_1\rho_1}{(1-\pi_1)\kappa_1\rho_0}-1},$$

$$A_1(1)=\frac{\frac{1-\frac{1}{\psi}}{1-\pi_1\kappa_1\rho_1}}{1-\frac{1-(\pi_0+\pi_1-1)\kappa_1\rho_1}{1-(\pi_0+\pi_1-1)\kappa_1\rho_0}\frac{(1-\pi_1)\kappa_1\rho_0}{1-\pi_1\kappa_1\rho_1}}.$$

Similarly, setting the coefficient of p_t to zero in (30) gives,

$$\begin{aligned}0 \;=\;& (-\theta A_0(0)+\theta A_0(1))+(\pi_0+\pi_1-1)\left\{\theta\kappa_1 A_0(0)+\frac{1}{2}\left[\left(-\frac{\theta}{\psi}+\theta\right)^2\sigma_0^2+\left(\theta\kappa_1 A_1(0)\varphi_e\right)^2\sigma_0^2\right]\right\}\\ & -(\pi_0+\pi_1-1)\left\{\theta\kappa_1 A_0(1)+\frac{1}{2}\left[\left(-\frac{\theta}{\psi}+\theta\right)^2\sigma_1^2+\left(\theta\kappa_1 A_1(1)\varphi_e\right)^2\sigma_1^2\right]\right\}.\end{aligned}$$

This implies,

$$A_0(0)=A_0(1)+\frac{C_1}{C_2}, \tag{33}$$

where $C_1=\frac{1}{2}(\pi_0+\pi_1-1)\left[\left(-\frac{\theta}{\psi}+\theta\right)^2(\sigma_0^2-\sigma_1^2)+\left(\theta\kappa_1 A_1(0)\varphi_e\right)^2\sigma_0^2-\left(\theta\kappa_1 A_1(1)\varphi_e\right)^2\sigma_1^2\right]$ and $C_2=\theta\left[1-(\pi_0+\pi_1-1)\kappa_1\right]$.

Finally, setting the constant in (30) to zero gives,

$$\begin{aligned}0 \;=\;& \theta\log\delta+\left(-\frac{\theta}{\psi}+\theta\right)\mu+\theta\kappa_0-\theta A_0(1)\\ & +(1-\pi_1)\left\{\theta\kappa_1 A_0(0)+\frac{1}{2}\left[\left(-\frac{\theta}{\psi}+\theta\right)^2\sigma_0^2+\left(\theta\kappa_1 A_1(0)\varphi_e\right)^2\sigma_0^2\right]\right\}\\ & +\pi_1\left\{\theta\kappa_1 A_0(1)+\frac{1}{2}\left[\left(-\frac{\theta}{\psi}+\theta\right)^2\sigma_1^2+\left(\theta\kappa_1 A_1(1)\varphi_e\right)^2\sigma_1^2\right]\right\}.\end{aligned}$$

Hence,

$$A_0(1) = \frac{(1-\pi_1)\,\kappa_1}{1-\pi_1\kappa_1} A_0(0) + \frac{C_3}{C_4}, \tag{34}$$

where, $C_3 = \theta \log \delta + \left(-\frac{\theta}{\psi} + \theta\right)\mu + \theta\kappa_0 + \frac{1}{2}(1-\pi_1)\left[\left(-\frac{\theta}{\psi}+\theta\right)^2 \sigma_0^2 + \left(\theta\kappa_1 A_1(0)\varphi_e\right)^2 \sigma_0^2\right] + \frac{1}{2}\pi_1\left[\left(-\frac{\theta}{\psi}+\theta\right)^2 \sigma_1^2 + \left(\theta\kappa_1 A_1(1)\varphi_e\right)^2 \sigma_1^2\right]$ and $C_4 = \theta\,(1-\pi_1\kappa_1)$.

Solving equations (33) and (34) for $A_0(0)$ and $A_0(1)$ gives,

$$A_0(0) = \frac{1-\pi_1\kappa_1}{1-\kappa_1}\left(\frac{C_1}{C_2} + \frac{C_3}{C_4}\right),$$

$$A_0(1) = \frac{(1-\pi_1)\,\kappa_1}{1-\kappa_1}\frac{C_1}{C_2} + \frac{1-\pi_1\kappa_1}{1-\kappa_1}\frac{C_3}{C_4}.$$

A.2 Dividend Claim

The market portfolio is defined as the claim to the aggregate dividend stream. Using the log-linear approximation for the continuous return on the aggregate dividend claim, $r_{m,t+1}$,

$$r_{m,t+1} = \kappa_{0,m} + \kappa_{1,m} z_{m,t+1} - z_{m,t} + \Delta c_{t+1},$$

where $z_{m,t}$ is the market-wide log price-dividend ratio. We conjecture that the log price-dividend ratio at date t takes the form,

$$z_{m,t} = p_t\left[A_{0,m}(0) + A_{1,m}(0) x_t\right] + (1-p_t)\left[A_{0,m}(1) + A_{1,m}(1) x_t\right].$$

The Euler equation for the dividend claim is,

$$E\left[\exp\left(m_{t+1} + r_{m,t+1}\right)|\mathcal{F}(t)\right] = 1. \tag{35}$$

Substituting the expression for m_{t+1} from (9) into (35), we have,

$$E\left[\exp\left(\theta\log\delta - \frac{\theta}{\psi}\Delta c_{t+1} + (\theta-1) r_{c,t+1} + r_{m,t+1}\right)|\mathcal{F}(t)\right] = 1. \tag{36}$$

Using the law of iterated expectations, equation (36) implies,

$$E\left[E\left\{\exp\left(\theta\log\delta - \frac{\theta}{\psi}\Delta c_{t+1} + (\theta-1) r_{c,t+1} + r_{m,t+1}\right)|\mathcal{F}(t), s_{t+1}\right\}|\mathcal{F}(t)\right] = 1. \tag{37}$$

Hence, using the conditional distribution of s_{t+1}, the above expression may be written as,

$$\sum_{i=0}^{1} P\left(s_{t+1}=i|\digamma(t)\right) E\left\{\exp\left(\theta\log\delta-\frac{\theta}{\psi}\Delta c_{t+1}+(\theta-1)r_{c,t+1}+r_{m,t+1}\right)|\digamma(t),s_{t+1}=i\right\}=1 \tag{38}$$

Proceeding as in Appendix A.1, equations (1), (4), (10), (11), (12), and (13) imply that,

$$\begin{aligned}
0 \;=\;& \left[\pi_0 p_t+\left(1-\pi_1\right)\left(1-p_t\right)\right]\times\{\theta\log\delta+\left(-\frac{\theta}{\psi}+\theta-1\right)\mu+\left(\theta-1\right)\kappa_0 \\
&+\left(\theta-1\right)\kappa_1 A_0(0)-\left(\theta-1\right)A_0(1)+\kappa_{0,m}+\kappa_{1,m}A_{0,m}(0)-A_{0,m}(1)+\mu_d \\
&+\left(-\frac{\theta}{\psi}+\left(\theta-1\right)\kappa_1 A_1(0)\rho_0-\left(\theta-1\right)A_1(1)+\theta-1+\kappa_{1,m}A_{1,m}(0)\rho_0-A_{1,m}(1)+\phi\right)x_t \\
&+\left(\left(\theta-1\right)\left[A_0(1)-A_0(0)\right]+A_{0,m}(1)-A_{0,m}(0)\right)p_t \\
&+\left(\left(\theta-1\right)\left[A_1(1)-A_1(0)\right]+A_{1,m}(1)-A_{1,m}(0)\right)p_t x_t \\
&+\frac{1}{2}\left[\left(-\frac{\theta}{\psi}+\theta-1\right)^2\sigma_0^2+\left((\theta-1)\kappa_1 A_1(0)+\kappa_{1,m}A_{1,m}(0)\right)^2\varphi_e^2\sigma_0^2+\varphi_d^2\sigma_0^2\right]\} \\
&+\left[1-\pi_0 p_t-\left(1-\pi_1\right)\left(1-p_t\right)\right]\times\{\theta\log\delta\left(-\frac{\theta}{\psi}+\theta-1\right)\mu+\left(\theta-1\right)\kappa_0 \\
&+\left(\theta-1\right)\kappa_1 A_0(1)-\left(\theta-1\right)A_0(1)+\kappa_{0,m}+\kappa_{1,m}A_{0,m}(1)-A_{0,m}(1)+\mu_d \\
&+\left(-\frac{\theta}{\psi}+\left(\theta-1\right)\kappa_1 A_1(1)\rho_1-\left(\theta-1\right)A_1(1)+\theta-1+\kappa_{1,m}A_{1,m}(1)\rho_1-A_{1,m}(1)+\phi\right)x_t \\
&+\left(\left(\theta-1\right)\left[A_0(1)-A_0(0)\right]+A_{0,m}(1)-A_{0,m}(0)\right)p_t \\
&+\left(\left(\theta-1\right)\left[A_1(1)-A_1(0)\right]+A_{1,m}(1)-A_{1,m}(0)\right)p_t x_t \\
&+\frac{1}{2}\left[\left(-\frac{\theta}{\psi}+\theta-1\right)^2\sigma_1^2+\left((\theta-1)\kappa_1 A_1(1)+\kappa_{1,m}A_{1,m}(1)\right)^2\varphi_e^2\sigma_1^2+\varphi_d^2\sigma_1^2\right]\}
\end{aligned}$$

The right-hand-side of the above expression may be simplified to,

$$
\begin{aligned}
&\left\{(\theta-1)\left[A_0(1)-A_0(0)\right]+A_{0,m}(1)-A_{0,m}(0)\right\}p_t\\
&+\left\{(\theta-1)\left[A_1(1)-A_1(0)\right]+A_{1,m}(1)-A_{1,m}(0)\right\}p_t x_t\\
&+\left(-\frac{\theta}{\psi}-(\theta-1)A_1(1)+\theta-1-A_{1,m}(1)+\phi\right)x_t\\
&+\left\{\theta\log\delta+\left(-\frac{\theta}{\psi}+\theta-1\right)\mu+(\theta-1)\kappa_0-(\theta-1)A_0(1)+\kappa_{0,m}-A_{0,m}(1)+\mu_d\right\}\\
&+\frac{1}{2}\left[(1-\pi_1)+(\pi_0+\pi_1-1)p_t\right]\left[\left(-\frac{\theta}{\psi}+\theta-1\right)^2\sigma_0^2+\left((\theta-1)\kappa_1A_1(0)+\kappa_{1,m}A_{1,m}(0)\right)^2\varphi_e^2\sigma_0^2+\varphi_d^2\sigma_0^2\right]\\
&+\frac{1}{2}\left[\pi_1-(\pi_0+\pi_1-1)p_t\right]\left[\left(-\frac{\theta}{\psi}+\theta-1\right)^2\sigma_1^2+\left((\theta-1)\kappa_1A_1(1)+\kappa_{1,m}A_{1,m}(1)\right)^2\varphi_e^2\sigma_1^2+\varphi_d^2\sigma_1^2\right]\\
&+\left[(1-\pi_1)+(\pi_0+\pi_1-1)p_t\right]\left((\theta-1)\kappa_1A_0(0)+\kappa_{1,m}A_{0,m}(0)+\left[(\theta-1)\kappa_1A_1(0)\rho_0+\kappa_{1,m}A_{1,m}(0)\rho_0\right]x\right.\\
&+\left[\pi_1-(\pi_0+\pi_1-1)p_t\right]\left((\theta-1)\kappa_1A_0(1)+\kappa_{1,m}A_{0,m}(1)+\left[(\theta-1)\kappa_1A_1(1)\rho_1+\kappa_{1,m}A_{1,m}(1)\rho_1\right]x_t\right)\quad\text{(3}
\end{aligned}
$$

Setting the coefficient of $p_t x_t$ to zero in the above equation gives,

$$
\begin{aligned}
0 \;=\;& (\theta-1)\left[A_1(1)-A_1(0)\right]+A_{1,m}(1)-A_{1,m}(0)\\
&+(\pi_0+\pi_1-1)\left\{(\theta-1)\kappa_1A_1(0)\rho_0+\kappa_{1,m}A_{1,m}(0)\rho_0-(\theta-1)\kappa_1A_1(1)\rho_1-\kappa_{1,m}A_{1,m}(1)\rho_1\right\},
\end{aligned}
$$

which implies,

$$
A_{1,m}(0)=\frac{D_1}{D_2}+\frac{D_3}{D_2}A_{1,m}(1) \tag{40}
$$

where $D_1=\left[1-(\pi_0+\pi_1-1)\kappa_1\rho_0\right]A_1(0)-\left[1-(\pi_0+\pi_1-1)\kappa_1\rho_1\right]A_1(1)$, $D_2=1-(\pi_0+\pi_1-1)\kappa_{1,m}\rho_0$, and $D_3=1-(\pi_0+\pi_1-1)\kappa_{1,m}\rho_1$.

Similarly, the coefficient on x_t to zero gives,

$$
A_{1,m}(1)=\frac{E_1}{E_2}+\frac{E_3}{E_2}A_{1,m}(0), \tag{41}
$$

where $E_1=\phi-1+(1-\pi_1\kappa_1\rho_1)A_1(1)-(1-\pi_1)\kappa_1\rho_0A_1(0)$, $E_2=1-\pi_1\kappa_{1,m}\rho_1$, and $E_3=(1-\pi_1)\kappa_{1,m}\rho_0$.

Solving equation (40) and (41) for $A_{1,m}(0)$ and $A_{1,m}(1)$ gives,

$$
A_{1,m}(0)=\frac{E_2D_1+E_1D_3}{E_2D_2-E_3D_3}
$$

$$
A_{1,m}(1)=\frac{E_1D_2+E_3D_1}{E_2D_2-E_3D_3}
$$

Similarly, setting the coefficient on p_t and the constant in (39) to zero and solving for $A_{0,m}(0)$ and $A_{0,m}(1)$ gives,

$$A_{0,m}(0) = \frac{F_2 G_1 + F_1 G_2}{F_2 \left(G_2 - G_3\right)}$$

$$A_{0,m}(1) = \frac{F_2 G_1 + F_1 G_3}{F_2 \left(G_2 - G_3\right)}$$

where $F_1 = (\theta - 1)\left[A_0(1) - A_0(0)\right]\left[1 - (\pi_0 + \pi_1 - 1)\kappa_1\right] +$
$\frac{1}{2}(\pi_0+\pi_1-1)\left[\left(-\frac{\theta}{\psi}+\theta-1\right)^2\sigma_0^2 + \left((\theta-1)\kappa_1 A_1(0) + \kappa_{1,m}A_{1,m}(0)\right)^2\varphi_e^2\sigma_0^2 + \varphi_d^2\sigma_0^2\right] -$
$\frac{1}{2}(\pi_0+\pi_1-1)\left[\left(-\frac{\theta}{\psi}+\theta-1\right)^2\sigma_1^2 + \left((\theta-1)\kappa_1 A_1(1) + \kappa_{1,m}A_{1,m}(1)\right)^2\varphi_e^2\sigma_1^2 + \varphi_d^2\sigma_1^2\right]$,
and $F_2 = 1 - (\pi_0 + \pi_1 - 1)\kappa_{1,m}$.

$G_1 = \theta\log\delta + \left(-\frac{\theta}{\psi}+\theta-1\right)\mu + (\theta-1)\kappa_0 + \kappa_{0,m} + \mu_d - (\theta-1)(1-\pi_1\kappa_1)A_0(1) +$
$(\theta-1)(1-\pi_1)\kappa_1 A_0(0)$
$+\frac{1}{2}(1-\pi_1)\left[\left(-\frac{\theta}{\psi}+\theta-1\right)^2\sigma_0^2 + \left((\theta-1)\kappa_1 A_1(0) + \kappa_{1,m}A_{1,m}(0)\right)^2\varphi_e^2\sigma_0^2 + \varphi_d^2\sigma_0^2\right]$
$+\frac{1}{2}\pi_1\left[\left(-\frac{\theta}{\psi}+\theta-1\right)^2\sigma_1^2 + \left((\theta-1)\kappa_1 A_1(1) + \kappa_{1,m}A_{1,m}(1)\right)^2\varphi_e^2\sigma_1^2 + \varphi_d^2\sigma_1^2\right]$,
$G_2 = 1 - \pi_1\kappa_{1,m}$, and $G_3 = (1-\pi_1)\kappa_{1,m}$.

A.3 Riskfree Rate

The risk free rate, $r_{f,t}$, is priced using the Euler equation,

$$E\left[\exp(m_{t+1} + r_{f,t}) | x_t, p_t\right] = 1.$$

Hence,

$$\begin{aligned} \frac{1}{R_{f,t}} &= E\left[\exp(m_{t+1}) | x_t, p_t\right] \\ &= E\left[\exp(\theta\log\delta - \frac{\theta}{\psi}\Delta c_{t+1} + (\theta-1)r_{c,t+1}) | \mathcal{F}(t)\right] \end{aligned}$$

As in the previous two subsections, the law of iterated expectations implies,

$$\frac{1}{R_{f,t}} = \sum_{i=0}^{1} P\left(s_{t+1} = i | \mathcal{F}(t)\right) E\left[\exp(\theta\log\delta - \frac{\theta}{\psi}\Delta c_{t+1} + (\theta-1)r_{c,t+1}) | \mathcal{F}(t), s_{t+1} = i\right] \tag{42}$$

Proceeding as in the prevoius subsection, equation (42) may be written as,

$$
\begin{aligned}
\frac{1}{R_{f,t}} = {} & \left[\pi_0 p_t + (1-\pi_1)(1-p_t)\right] \times \Big(\Big\{\theta \log\delta + \left(-\frac{\theta}{\psi} + \theta - 1\right)\mu + (\theta-1)\kappa_0 + (\theta-1)\kappa_1 A_0(0) \\
& -(\theta-1)A_0(1) + \left(-\frac{\theta}{\psi} + (\theta-1)\kappa_1 A_1(0)\rho_0 - (\theta-1)A_1(1) + \theta - 1\right) x_t \\
& +(\theta-1)\left(A_0(1) - A_0(0)\right) p_t + (\theta-1)\left(A_1(1) - A_1(0)\right) p_t x_t \\
& +\frac{1}{2}\left[\left(-\frac{\theta}{\psi} + \theta - 1\right)^2 \sigma_0^2 + \left((\theta-1)\kappa_1 A_1(0)\varphi_e\right)^2 \sigma_0^2\right]\Big\} + 1\Big) \\
& + \left[1 - \pi_0 p_t - (1-\pi_1)(1-p_t)\right] \times \Big(\Big\{\theta \log\delta + \left(-\frac{\theta}{\psi} + \theta - 1\right)\mu + (\theta-1)\kappa_0 + (\theta-1)\kappa_1 A_0(1 \\
& -(\theta-1)A_0(1) + \left(-\frac{\theta}{\psi} + (\theta-1)\kappa_1 A_1(1)\rho_1 - (\theta-1)A_1(1) + \theta - 1\right) x_t \\
& +(\theta-1)\left(A_0(1) - A_0(0)\right) p_t + (\theta-1)\left(A_1(1) - A_1(0)\right) p_t x_t \\
& +\frac{1}{2}\left[\left(-\frac{\theta}{\psi} + \theta - 1\right)^2 \sigma_1^2 + \left((\theta-1)\kappa_1 A_1(1)\varphi_e\right)^2 \sigma_1^2\right]\Big\} + 1\Big)
\end{aligned}
$$

Simplifying the above expression gives,

$$
\begin{aligned}
\frac{1}{R_{f,t}} = {} & \theta \log\delta + \left(-\frac{\theta}{\psi} + \theta - 1\right)\mu + (\theta-1)\kappa_0 - (\theta-1)A_0(1) + \left(-\frac{\theta}{\psi} - (\theta-1)A_1(1) + \theta - 1\right) x_t \\
& +(\theta-1)\left(A_0(1) - A_0(0)\right) p_t + (\theta-1)\left(A_1(1) - A_1(0)\right) p_t x_t \\
& + \left[(1-\pi_1) + (\pi_0 + \pi_1 - 1)p_t\right]\left[(\theta-1)\kappa_1 A_0(0) + (\theta-1)\kappa_1 A_1(0)\rho_0 x_t\right] \\
& +\frac{1}{2}\left[(1-\pi_1) + (\pi_0 + \pi_1 - 1)p_t\right]\left[\left(-\frac{\theta}{\psi} + \theta - 1\right)^2 \sigma_0^2 + \left((\theta-1)\kappa_1 A_1(0)\varphi_e\right)^2 \sigma_0^2\right] \\
& + \left[\pi_1 + (1 - \pi_0 - \pi_1)p_t\right]\left[(\theta-1)\kappa_1 A_0(1) + (\theta-1)\kappa_1 A_1(1)\rho_1 x_t\right] \\
& +\frac{1}{2}\left[\pi_1 + (1 - \pi_0 - \pi_1)p_t\right]\left[\left(-\frac{\theta}{\psi} + \theta - 1\right)^2 \sigma_1^2 + \left((\theta-1)\kappa_1 A_1(1)\varphi_e\right)^2 \sigma_1^2\right] + 1
\end{aligned}
$$

$$
\begin{aligned}
= \quad & \theta \log \delta + \left(-\frac{\theta}{\psi} + \theta - 1\right)\mu + (\theta-1)\kappa_0 - (\theta-1)A_0(1) + (1-\pi_1)\,(\theta-1)\kappa_1 A_0(0) \\
& +\pi_1(\theta-1)\kappa_1 A_0(1) + \frac{1}{2}\,(1-\pi_1)\left[\left(-\frac{\theta}{\psi} + \theta - 1\right)^2 \sigma_0^2 + \left((\theta-1)\kappa_1 A_1(0)\varphi_e\right)^2 \sigma_0^2\right] \\
& +\frac{1}{2}\pi_1\left[\left(-\frac{\theta}{\psi} + \theta - 1\right)^2 \sigma_1^2 + \left((\theta-1)\kappa_1 A_1(1)\varphi_e\right)^2 \sigma_1^2\right] + 1 \\
& +\left(-\frac{\theta}{\psi} - (\theta-1)A_1(1) + \theta - 1 + (1-\pi_1)\,(\theta-1)\kappa_1 A_1(0)\rho_0 + \pi_1(\theta-1)\kappa_1 A_1(1)\rho_1\right) x_t \\
& +[(\theta-1)\,(A_0(1) - A_0(0)) \\
& +(\pi_0+\pi_1-1)\left\{(\theta-1)\kappa_1 A_0(0) + \frac{1}{2}\left[\left(-\frac{\theta}{\psi} + \theta - 1\right)^2 \sigma_0^2 + \left((\theta-1)\kappa_1 A_1(0)\varphi_e\right)^2 \sigma_0^2\right]\right\} \\
& +(1-\pi_0-\pi_1)\left\{(\theta-1)\kappa_1 A_0(1) + \frac{1}{2}\left[\left(-\frac{\theta}{\psi} + \theta - 1\right)^2 \sigma_1^2 + \left((\theta-1)\kappa_1 A_1(1)\varphi_e\right)^2 \sigma_1^2\right]\right\}]p_t \\
& +[(\theta-1)\,(A_1(1) - A_1(0)) + (\pi_0+\pi_1-1)(\theta-1)\kappa_1 A_1(0)\rho_0 \\
& +(1-\pi_0-\pi_1)(\theta-1)\kappa_1 A_1(1)\rho_1]x_t p_t
\end{aligned}
$$

Therefore,

$$
\frac{1}{R_{f,t}} = A_{0,f} + A_{1,f}x_t + A_{2,f}p_t + A_{3,f}x_t p_t
$$

where

$$\begin{aligned}
A_{0,f} &= \theta\log\delta + \left(-\frac{\theta}{\psi}+\theta-1\right)\mu + (\theta-1)\kappa_0 - (\theta-1)A_0(1) + (1-\pi_1)(\theta-1)\kappa_1 A_0(0) \\
&\quad +\pi_1(\theta-1)\kappa_1 A_0(1) + \frac{1}{2}(1-\pi_1)\left[\left(-\frac{\theta}{\psi}+\theta-1\right)^2\sigma_0^2 + \left((\theta-1)\kappa_1 A_1(0)\varphi_e\right)^2\sigma_0^2\right] \\
&\quad +\frac{1}{2}\pi_1\left[\left(-\frac{\theta}{\psi}+\theta-1\right)^2\sigma_1^2 + \left((\theta-1)\kappa_1 A_1(1)\varphi_e\right)^2\sigma_1^2\right] + 1 \\
A_{1,f} &= -\frac{\theta}{\psi} - (\theta-1)A_1(1) + \theta - 1 + (1-\pi_1)(\theta-1)\kappa_1 A_1(0)\rho_0 + \pi_1(\theta-1)\kappa_1 A_1(1)\rho_1 \\
A_{2,f} &= (\pi_0+\pi_1-1)\left\{(\theta-1)\kappa_1 A_0(0) + \frac{1}{2}\left[\left(-\frac{\theta}{\psi}+\theta-1\right)^2\sigma_0^2 + \left((\theta-1)\kappa_1 A_1(0)\varphi_e\right)^2\sigma_0^2\right]\right\} \\
&\quad +(1-\pi_0-\pi_1)\left\{(\theta-1)\kappa_1 A_0(1) + \frac{1}{2}\left[\left(-\frac{\theta}{\psi}+\theta-1\right)^2\sigma_1^2 + \left((\theta-1)\kappa_1 A_1(1)\varphi_e\right)^2\sigma_1^2\right]\right\} \\
&\quad +(\theta-1)\left(A_0(1)-A_0(0)\right) \\
A_{3,f} &= (\theta-1)\left(A_1(1)-A_1(0)\right) + (\pi_0+\pi_1-1)(\theta-1)\kappa_1\left(A_1(0)\rho_0 - A_1(1)\rho_1\right)
\end{aligned}$$

A.4 Equity Premium

Using the log-linearized return on the market portfolio in equation (11) and noting that the log price-dividend ratio of the market is given by equation (13), we have

$$\begin{aligned}
r_{m,t+1} &= \kappa_{0,m} + \kappa_{1,m}\left(p_{t+1}\left[A_{0,m}(0) + A_{1,m}(0)x_{t+1}\right] + (1-p_{t+1})\left[A_{0,m}(1) + A_{1,m}(1)x_{t+1}\right]\right) \\
&\quad - \left(p_t\left[A_{0,m}(0) + A_{1,m}(0)x_t\right] + (1-p_t)\left[A_{0,m}(1) + A_{1,m}(1)x_t\right]\right) \\
&\quad +\mu_d + \phi x_t + \varphi_d\sigma_{s_{t+1}}u_{t+1}. \qquad (43)
\end{aligned}$$

Taking conditional expectations of both sides of equation (43) with respect to the time t information set, $\mathcal{F}(t)$, and noting that:

$$\begin{aligned}
E\left(\sigma_{s_{t+1}}u_{t+1}|\mathcal{F}(t)\right) &= 0 \\
E\left(x_{t+1}|\mathcal{F}(t)\right) &= \left\{\left[(1-\pi_1) + (\pi_0+\pi_1-1)p_t\right]\rho_0 + \left[\pi_1 - (\pi_0+\pi_1-1)p_t\right]\rho_1\right\}x_t \\
E\left(p_{t+1}|\mathcal{F}(t)\right) &= (1-\pi_1) + (\pi_0+\pi_1-1)p_t \\
E\left(p_{t+1}x_{t+1}|\mathcal{F}(t)\right) &= (1-\pi_1)\left\{\left[(1-\pi_1) + (\pi_0+\pi_1-1)p_t\right]\rho_0 + \left[\pi_1 - (\pi_0+\pi_1-1)p_t\right]\rho_1\right\}x_t \\
&\quad (\pi_0+\pi_1-1)\left\{\left[(1-\pi_1) + (\pi_0+\pi_1-1)p_t\right]\rho_0 + \left[\pi_1 - (\pi_0+\pi_1-1)p_t\right]\rho_1\right\}p_t x_t
\end{aligned}$$

we have

$$E\left(r_{m,t+1}|\mathcal{F}(t)\right) = B_0 + B_1 x_t + B_2 p_t + B_3 p_t x_t + B_4 p_t^2 x_t$$

where

$$\begin{array}{rcl} B_0 & = & \kappa_{0,m} + (\kappa_{1,m} - 1)A_{0,m}(1) + \kappa_{1,m}\left[A_{0,m}(0) - A_{0,m}(1)\right](1-\pi_1) + \mu_d \\ B_1 & = & \kappa_{1,m}A_{1,m}(1)\left[(1-\pi_1)\rho_0 + \pi_1\rho_1\right] - A_{1,m}(1) + \phi \\ & & +(1-\pi_1)\left[(1-\pi_1)\rho_0 + \pi_1\rho_1\right]\kappa_{1,m}\left[A_{1,m}(0) - A_{1,m}(1)\right] \\ B_2 & = & \left[A_{0,m}(0) - A_{0,m}(1)\right]\left((\pi_0 + \pi_1 - 1)\kappa_{1,m} - 1\right) \\ B_3 & = & \left[A_{1,m}(0) - A_{1,m}(1)\right](\pi_0 + \pi_1 - 1)\kappa_{1,m}\left[(1-\pi_1)(\rho_0 - \rho_1) + (1-\pi_1)\rho_0 + \pi_1\rho_1 - 1\right] \\ & & +\kappa_{1,m}A_{1,m}(1)(\pi_0 + \pi_1 - 1)(\rho_0 - \rho_1) \\ B_4 & = & \left[A_{1,m}(0) - A_{1,m}(1)\right](\pi_0 + \pi_1 - 1)^2(\rho_0 - \rho_1)\kappa_{1,m} \end{array}$$

Now, the gross risk free rate is given by equation (14)

$$\frac{1}{R_{f,t}} = A_{0,f} + A_{1,f}x_t + A_{2,f}p_t + A_{3,f}x_t p_t.$$

Hence, the equity premium is given by

$$E\left[\left(r_{m,t+1} - \frac{1}{R_{f,t}}\right)|\mathcal{F}(t)\right] = E_0 + E_1 x_t + E_2 p_t + E_3 p_t x_t + E_4 p_t^2 x_t$$

where $E_i = B_i - A_{i,f}$, $i = 0, 1, 2, 3$, and $E_4 = B_4$.

A.5 Pricing Kernel

The pricing kernel is given by equation (9),

$$m_{t+1} = \theta \log \delta - \frac{\theta}{\psi}\Delta c_{t+1} + (\theta - 1)r_{c,t+1}.$$

Now, the log-linearization in equation (10),

$$r_{c,t+1} = \kappa_0 + \kappa_1 z_{t+1} - z_t + \Delta c_{t+1},$$

and the conjecture (12), that was verified in Subsection A.1,

$$z_t = p_t\left[A_0(0) + A_1(0)x_t\right] + (1-p_t)\left[A_0(1) + A_1(1)x_t\right],$$

together imply that,

$$\begin{aligned} m_{t+1} &= \theta \log \delta - \frac{\theta}{\psi} \Delta c_{t+1} + (\theta - 1)\kappa_0 \\ &+ (\theta-1)\kappa_1 p_{t+1} \left[A_0(0) + A_1(0) x_{t+1} \right] \\ &+ (\theta-1)\kappa_1 (1 - p_{t+1}) \left[A_0(1) + A_1(1) x_{t+1} \right] \\ &- (\theta-1) p_t \left[A_0(0) + A_1(0) x_t \right] - (\theta-1)(1-p_t) \left[A_0(1) + A_1(1) x_t \right] \\ &+ (\theta-1)\Delta c_{t+1}. \end{aligned}$$

Collecting terms in the above expression, we have,

$$m_{t+1} = c_0 + c_1 \Delta c_{t+1} + c_2 p_{t+1} + c_3 p_t + c_4 x_{t+1} + c_5 x_t + c_6 p_{t+1} x_{t+1} + c_7 p_t x_t,$$

where,

$$\begin{aligned} c_0 &= \theta \log(\delta) + (\theta-1)\kappa_0 + (\theta-1)(\kappa_1 - 1) A_0(1), \\ c_1 &= -\frac{\theta}{\psi} + \theta - 1, \\ c_2 &= (\theta-1)\kappa_1 \left[A_0(0) - A_0(1) \right], \\ c_3 &= -(\theta-1) \left[A_0(0) - A_0(1) \right], \\ c_4 &= (\theta-1)\kappa_1 A_1(1), \\ c_5 &= -(\theta-1) A_1(1), \\ c_6 &= (\theta-1)\kappa_1 \left[A_1(0) - A_1(1) \right], \\ c_7 &= -(\theta-1) \left[A_1(0) - A_1(1) \right]. \end{aligned}$$

This gives equation (16) in the paper. Note that this expression for the pricing kernel involves the state variables, x_t and p_t. These are latent to the econometrician. However, note that it was shown in Subsections A.2 and A.3, respectively, that the log price-dividend ratio of the aggregate stock market, $z_{m,t}$, and the gross risk free rate, $R_{f,t}$, are functions only of these two latent state variables (equations (13) and (14)),

$$\begin{aligned} z_{m,t} &= p_t \left[A_{0,m}(0) + A_{1,m}(0) x_t \right] + (1-p_t) \left[A_{0,m}(1) + A_{1,m}(1) x_t \right] \\ &= A_{0,m}(1) + A_{1,m}(1) x_t + \left[A_{0,m}(0) - A_{0,m}(1) \right] p_t + \left[A_{1,m}(0) - A_{1,m}(1) \right] p_t x_t, \end{aligned}$$

$$\frac{1}{R_{f,t}} = A_{0,f} + A_{1,f} x_t + A_{2,f} p_t + A_{3,f} x_t p_t.$$

Therefore, the above two equations may be inverted to express the latent state variables, x_t and p_t, as functions of the observables, $z_{m,t}$ and $R_{f,t}$. In particular, (13) implies

$$x_t = \frac{\frac{1}{R_{f,t}} - A_{0,f} - A_{2,f} p_t}{A_{1,f} + A_{3,f} p_t}. \tag{44}$$

Substituting (44) into (14), and simplifying gives the following quadratic equation for p_t:

$$a p_t^2 + b_t p_t + h_t = 0, \tag{45}$$

where

$$\begin{aligned}
a &= A_{3,f}\left[A_{0,m}(0) - A_{0,m}(1)\right] - A_{2,f}\left[A_{1,m}(0) - A_{1,m}(1)\right], \\
b_t &= \left[A_{1,m}(0) - A_{1,m}(1)\right]\left(\frac{1}{R_{f,t}} - A_{0,f}\right) + A_{1,f}\left[A_{0,m}(0) - A_{0,m}(1)\right] \\
&\quad + A_{0,m}(1) A_{3,f} - A_{1,m}(1) A_{2,f} - z_{m,t} A_{3,f}, \\
h_t &= A_{1,m}(1)\left(\frac{1}{R_{f,t}} - A_{0,f}\right) + A_{0,m}(1) A_{1,f} - z_{m,t} A_{1,f}.
\end{aligned}$$

Equation (45) implies two solutions for p_t in terms of the observables,$z_{m,t}$ and $R_{f,t}$, given by

$$p_t = \frac{-b_t \pm \sqrt{b_t^2 - 4 a h_t}}{2a} \tag{46}$$

Substituting the solutions in (46) into (44) gives the two corresponding solutions for x_t in terms of the observables,$z_{m,t}$ and $R_{f,t}$.

A.6 Simulated Moments Estimation Methodology

The Simulated Moments Estimation (SME) methodology of Duffie and Singleton (1993) extends the GMM approach of Hansen (1982) to models for which the moment restrictions of interest do not have analytic representations in terms of observable variables and the unknown parameter vector, β_0. In what follows, we present this econometric methodology and discuss the procedures for constructing standard errors and test statistics for overidentifying restrictions. We also present an extension of the methodology in Duffie and Singleton (1993) that accommodates the possibility that a subset of the moment restrictions have analytic representations in terms of observable variables and the unknown parameter vector while the others do not. This extension is relevant for the estimation problem in this paper where the Euler equations for a set of assets deliver moment restrictions that are known functions of observables like the aggregate consumption growth rate, the market-wide price-dividend ratio, the risk free rate, and the returns on the set of assets the model is asked to price, and the unknown time-series and preference parameters, but the time-series moment restrictions including the means, variances and autocovariance functions of aggregate consumption and dividend growth rates do not have analytic representations.

Consider a model that delivers observables as $f(z_t;\beta_0)$, where f is an R^q-valued function, z_t is an l-dimensional vector of state variables, and β is a p-dimensional vector of parameters $(q \geqslant p)$ with compact parameter set Θ and true value β_0. The econometrician observes draws of an R^q-valued random variable, $f_t^* \equiv f(z_t;\beta)$. In the special case when the function mapping β to $E[f(z_t;\beta)]$ is known and independent of t, the parameter vector β may be estimated using the GMM approach of Hansen (1982). Specifically,

$$\widehat{\beta}_{GMM} = \underset{\beta\in\Theta}{arg\ min} \left[\frac{1}{T}\sum_{t=1}^{T} f_t^* - E[f(z_t;\beta)]\right]' W_T \left[\frac{1}{T}\sum_{t=1}^{T} f_t^* - E[f(z_t;\beta)]\right],$$

where W_T is a sequence of positive definite weighting matrices.

In a large class of asset pricing models, the mapping $\beta \mapsto E[f(z_t;\beta)]$ is not known. In particular, for the time-series specification of the regime shifts model presented in Section 2, moments of consumption and dividend growth rates do not have closed-form expressions in terms of the unknown parameter vector β.

However, apart from a measurable observation function f, another basic primitive for these models is a transition function H such that the state process $\{z_t\}_{t=1}^{\infty}$ is generated by the difference equation

$$z_{t+1} = H\left(z_t, \zeta_{t+1}, \beta_0\right),$$

where $\{\zeta_t\}$ is an *i.i.d.* sequence of random variables on a given probability space $(\Omega, \mathcal{F}, P)$. The simulated moments estimator, rather than requiring $E[f(z_t;\beta)]$ to be known in closed-form and independent of t, relies on the much weaker assumption that the econometrician has access to a sequence $\left\{\widehat{\zeta}_t\right\}$ of random variables that is identically and independently distributed of $\{\zeta_t\}$. Then, given any initial condition for the state vector $\widehat{z}_1$, and any admissible parameter vector $\beta \in \Theta$, a history $\left\{z_t^\beta\right\}_{t=1}^{N}$ of N simulated equilibrium states can be generated by letting $z_1^\beta = \widehat{z}_1$ and

$$z_{t+1}^\beta = H\left(z_t^\beta, \widehat{\zeta}_{t+1}, \beta_0\right).$$

Here, $N : \mathbb{N} \to \mathbb{N}$ denotes the simulation sample size $N(T)$ that is generated for a given sample size T of actual observations, where $N(T) \to \infty$ AS $T \to \infty$. Likewise, the simulated observation process $\left\{f_t^\beta\right\}_{t=1}^{N}$ is constructed by $f_t^\beta = f(z_t^\beta;\beta)$. The SME is a value of β chosen to minimize the distance between the sample mean of $\left\{ f_t^{\beta} \right\}_{t=1}^{N}$ and the sample mean of $\{f_t^*\}_{t=1}^{T}$, where T is the number of historical observations on f_t^*:

$$\widehat{\beta}_{SME} = \underset{\beta \in \Theta}{arg\ min} G_T(\beta)' W_T G_T(\beta)$$

where $G_T(\beta) = \left[\frac{1}{T} \sum_{t=1}^{T} f_t^* - \frac{1}{N} \sum_{s=1}^{N} f_s^{\beta} \right]$.

We first estimate the 11 time-series parameters using 13 time-series moment restrictions corresponding to the unconditional means, variances, and first, second, and third-order autocovariances of consumption and dividend growth rates, the covariance between consumption and dividend growth rates, and the covariance between consumption growth and one and two lags of the dividend growth rate. In this case, f_t^* does not depend on the parameter β. Under ceratin regularity conditions,[13] Duffie and Singleton (1993) show that if the weighting matrix W_T is chosen such that $W_T \rightarrow W_0 = \Sigma_0^{-1}$ almost surely, where (for any t)

$$\Sigma_0 \equiv \sum_{j=-\infty}^{\infty} E\left(\left[f_t^* - E\left(f_t^*\right) \right] \left[f_{t-j}^* - E\left(f_{t-j}^* \right) \right]' \right), \tag{47}$$

then $\sqrt{T}\left(\widehat{\beta}_{SME} - \beta_0 \right)$ converges in distribution as $T \rightarrow \infty$ to a normal random vector with mean zero and covariance matrix

$$\Lambda = (1+\tau)\left(D_0' \Sigma_0^{-1} D_0 \right)^{-1},$$

where $D_0 = E\left(\frac{\partial}{\partial \beta'} f_{\infty}^{\beta_0} \right)$ and $\frac{T}{N(T)} \rightarrow \tau$ an $T \rightarrow \infty$.

As with the GMM approach, an overidentifying restrictions test may be performed to test the specification of the model. Given the normalized asymptotic distribution of the estimator, the following statistic converges in distribution as $T \rightarrow \infty$ to a chisquared random variable with $q-p$ degrees of freedom under the null that the model is correctly specified:

$$\frac{T}{1+\tau} G_T(\widehat{\beta}_{SME})' \widehat{\Sigma}^{-1} G_T(\widehat{\beta}_{SME}) \xrightarrow{d} \chi^2_{q-p}.$$

We next estimate the time-series and preference parameters using simultaneously the pricing restrictions given by the Euler equations and the time-series moment restrictions implied by the time-series specification of the model. Note that the Euler equations for a set of assets deliver moment restrictions that are known functions of observables like the aggregate consumption growth rate, the market-wide price-dividend

---

[13] see Duffie and Singleton (1993) for details.

ratio, the risk free rate, and the returns on the set of assets, and the unknown time-series and preference parameters, but the time-series moment restrictions including the means, variances and autocovariance functions of aggregate consumption and dividend growth rates do not have analytic representations. In this case, we adopt a hybrid estimation methodology that combines features of the GMM and SME approaches.

The moment restrictions may be written as

$$G_T(\beta) = \begin{pmatrix} G_{1T}(\beta) \\ G_{2T}(\beta) \end{pmatrix} = \begin{pmatrix} \frac{1}{T}\sum_{t=1}^{T} f_{1t}^{\beta} \\ \frac{1}{T}\sum_{t=1}^{T} f_{2t}^{*} - \frac{1}{N}\sum_{s=1}^{N} f_{2s}^{\beta} \end{pmatrix},$$

where $G_{1T}(\beta)$ are the moment restrictions corresponding to the Euler equations, with typical element $\frac{1}{T}\sum_{t=1}^{T} exp\left(m_{t+1}(\beta) + r_{j,t+1}\right) - 1$. Note that these have analytic expressions in terms of observable variables and the unknown parameter vector β. Also, $E\left[f_{1t}^{\beta_0}\right] = 0$. $\frac{1}{T}\sum_{t=1}^{T} f_{2t}^{*}$ are the moment conditions corresponding to the time-series restrictions such that $E\left[f_{2t}^{*}\right]$ does not have an analytic representation in terms of the parameters. Hence, for this subset, the moment restrictions are expressed as the distance between the sample mean of the simulated sequence $\left\{f_{2t}^{\beta}\right\}_{t=1}^{N}$ and the sample mean of the historical data $\left\{f_{2t}^{*}\right\}_{t=1}^{T}$.

Hence, we have,

$$\sqrt{T}G_T(\beta) = \begin{pmatrix} \frac{1}{\sqrt{T}}\sum_{t=1}^{T} f_{1t}^{\beta} \\ \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[f_{2t}^{*} - E\left(f_{2\infty}^{*}\right)\right] - \frac{\sqrt{T}}{\sqrt{N}}\left(\frac{1}{\sqrt{N}}\sum_{s=1}^{N}\left[f_{2s}^{\beta} - E\left(f_{2\infty}^{\beta}\right)\right]\right) \end{pmatrix}. \tag{48}$$

Consider the first set of moment restrictions in (48). Under regularity conditions in Pakes and Pollard (1989), we have

$$\sqrt{T}G_{1T}(\beta_0) \xrightarrow{d} N(0, V_{11}), \tag{49}$$

where $V_{11,0} = E\left[f_{1t}^{\beta_0}\left(f_{1t}^{\beta_0}\right)'\right]$.

Analogously, under regularity conditions in Duffie and Singleton (1993), we have

$$\sqrt{T}G_{2T}(\beta_0) \xrightarrow{d} N(0, V_{22}), \tag{50}$$

where $V_{22,0} = (1+\tau)\,\Sigma_0$, where Σ_0 is defined in equation (47).

Finally, the asymptotic covariance between $\sqrt{T}G_{1T}(\beta_0)$ and $\sqrt{T}G_{2T}(\beta_0)$ can be shown to be

$$V_{12,0} = \sum_{j=-\infty}^{\infty} E\left(\left[f_{1t}^{\beta_0} - E\left(f_{1t}^{\beta_0}\right)\right]\left[f_{2,t-j}^{*} - E\left(f_{2,\infty}^{*}\right)\right]'\right). \tag{51}$$

Combining equations (49), (50), and (51) gives

$$\sqrt{T}G_T(\beta_0) \xrightarrow{d} N(0, V_0),$$

where $V_0 = \begin{pmatrix} V_{11,0} & V_{12,0} \\ V_{12,0}' & V_{22,0} \end{pmatrix}$.

Therefore, if the weighting matrix W_T is chosen such that $W_T \to W_0 = V_0^{-1}$ almost surely, then $\sqrt{T}\left(\widehat{\beta}_{SME} - \beta_0\right)$ converges in distribution as $T \to \infty$ to a normal random vector with mean zero and covariance matrix

$$\Lambda = \left(D_0' V_0^{-1} D_0\right)^{-1},$$

where $D_0 = E\begin{pmatrix} \frac{\partial}{\partial\beta'} f_{1t}^{\beta_0} \\ -\frac{\partial}{\partial\beta'} f_{2,\infty}^{\beta_0} \end{pmatrix}$.

An overidentifying restrictions test may be performed to test the specification of the model. Given the normalized asymptotic distribution of the estimator, the following statistic converges in distribution as $T \to \infty$ to a chisquared random variable with $q - p$ degrees of freedom under the null that the model is correctly specified:

$$G_T(\widehat{\beta}_{SME})' \widehat{V}^{-1} G_T(\widehat{\beta}_{SME}) \xrightarrow{d} \chi^2_{q-p}.$$

A.7 Simulation Design

We obtain the finite-sample distribution of the J-stat for the overidentifying restrictions with Monte Carlo simulation. We calibrate the parameters of the time series to their SME point estimates and set the initial conditions of the state variables to their unconditional means, $x_0 = 0$ and $p_0 = \frac{1-\pi_1}{2-\pi_0-\pi_1}$. Given the initial value, p_0, we simulate a time-series of p_t of the same size as the historical sample. We use this simulated time-series to draw the state of the economy, i.e. the regime, at each time period, and, hence, simulate the time-series of the LRR variable, the aggregate consumption and dividend growth rates. For the 2-asset system, we simulate the time-series of log returns on the market portfolio and the log risk free rate, using the log-linearization in equation (11) and the model solution in equation (14), respectively. For the 6-asset system, we simulate the series for the log returns on the Small, Large, Growth, and Value portfolios, using similar log-linearizations as for the market portfolio. We then perform the SME estimation of the time-series and preference parameters using jointly the pricing and the time-series restrictions for the two-asset and six-asset systems, as in the empirical section 5. We also compute the J-stat for the overidentifying restrictions. We repeat the simulation 100 times and obtain the 90%, 95%, and 99% critical values of the J-stat from its finite-sample distribution. We perform the simulation for the 2-asset and the 6-asset systems for the full-sample period 1930-2006.

Table 1: Descriptive Statistics

	$log(returns)$		$log(P/D)$		$log(D_{t+1}/D_t)$	
	$Mean$	$Std.Dev.$	$Mean$	$Std.Dev$	$Mean$	$Std.Dev$
$SizePortfolios$						
$Small$	0.105	0.333	4.147	0.711	0.083	0.347
$L\arg e$	0.060	0.184	3.289	0.440	0.012	0.136
$B/M\ Portfolios$						
$Growth$	0.052	0.206	3.725	0.630	0.007	0.206
$Value$	0.093	0.302	3.588	1.135	0.070	0.568
$Market$	0.066	0.193	3.267	0.384	0.014	0.108
$Risk\ free\ rate$	0.008	0.050				

This table reports the descriptive statistics for the annual log returns, the log price-dividend ratios, and the log dividend growth rates of the market, the risk free rate, the "Small", "Large", "Growth", and "Value" portfolios. The sample period is 1930-2006.

Table 2: Tests of the Time-Series Specification

	Panel A: identity weighting matrix										
	μ	μ_d	ϕ	φ_d	ρ_1	ρ_0	φ_e	σ_1	σ_0	π_1	π_0
$Time - Series$	0.015 (0.005)	0.014 (0.017)	4.0 (0.023)	6.0 (0.016)	0.92 (0.274)	0.50 (0.322)	0.3 (0.219)	0.012 (0.613)	0.021 (0.238)	0.70 (0.053)	0.80 (0.089)
$Dist - stat$	0.001 (>0.10)										
	Panel B: efficient weighting matrix										
	μ	μ_d	ϕ	φ_d	ρ_1	ρ_0	φ_e	σ_1	σ_0	π_1	π_0
$Time - Series$	0.015 (0.002)	0.014 (0.011)	3.0 (0.355)	4.0 (0.317)	0.92 (0.345)	0.50 (0.345)	0.5 (0.175)	0.012 (0.245)	0.021 (0.095)	0.80 (0.0352)	0.80 (0.020)
$J - stat$	8.528 (>0.10)										

The table reports the SME estimates of the time-series parameters and the test statistics for over-identifying restrictions using annual data over the period 1930-2006. Panels A and B report results for the identity and the efficient weighting matrices, respectively. 13 time-series moment restrictions on the unconditional moments of aggregate consumption and dividend growth rates are used in the estimation. The number of time-series parameters to be estimated is 11. The first row in each panel presents the parameter estimates along with the associated standard errors in parentheses. Standard errors are Newey-West corrected using 2 lags. The second row in each panel reports the test statistic for the over-identifying restrictions along with the associated finite-sample p-value in parentheses. For the identity weighting matrix, the statistic has a non-standard aymptotic distribution with 90%, 95%, and 99% critical values equal to 0.0003, 0.0005, and 0.0010, respectively. For the efficient weighting matrix, the statistic has an aymptotic chisquared distribution with 2 degrees of freedom. The 90%, 95%, and 99% critical values from the finite-sample distribution of the Dist-stat are 0.01, 0.02, and 0.05, respectively. The 90%, 95%, and 99% critical values from the finite-sample distribution of the J-stat are 30.23, 35.22, and 53.91, respectively.

Table 3: Tests on the 2-Asset System

Panel A: identity weighting matrix

	μ	μ_d	ϕ	φ_d	ρ_1	ρ_0	φ_e	σ_1	σ_0	π_1	π_0	δ	γ	ψ
Estimates	0.015 (0.004)	0.014 (0.015)	2.0 (0.085)	6.0 (0.005)	0.94 (0.088)	0.6 (0.164)	0.3 (0.111)	0.009 (0.118)	0.024 (0.008)	0.5 (0.079)	0.9 (0.068)	0.99 (0.031)	10 (0.012)	0.6 (0.139)
Pr*icing Errors*		*Mean*	*Std.Err.*											
Market		0.035	0.068											
Risk free rate		−0.029	0.063											
Dist − stat	1.373 (<0.01)													

Panel B: efficient weighting matrix

	μ	μ_d	ϕ	φ_d	ρ_1	ρ_0	φ_e	σ_1	σ_0	π_1	π_0	δ	γ	ψ
Estimates	0.015 (0.006)	0.014 (0.015)	4.0 (0.351)	4.0 (0.352)	0.96 (0.295)	0.7 (0.262)	0.5 (0.215)	0.009 (0.076)	0.015 (0.037)	0.5 (0.060)	0.6 (0.111)	0.99 (0.032)	6 (0.369)	0.6 (0.286)
Pr*icing Errors*		*Mean*	*Std.Err.*											
Market		0.081	0.053											
Risk free rate		0.001	0.045											
J − stat	24.53 (>0.10)													

The table reports SME estimates using annual data over the period 1930-2006. Panels A and B report results for the identity and the efficient weighting matrices, respectively. The asset menu consists of the market portfolio and the risk free rate. The lagged log price-dividend ratio of the market and the lagged log risk free rate are used as instruments, giving 6 pricing restrictions. To this set of pricing restrictions, we add the 13 moment restrictions implied by the time-series specification of the model. This gives a total of 19 moment conditions. The total number of parameters to be estimated is 14, including 11 time-series and 3 preference parameters. The bigger root of the quadratic is used in the extraction of the latent state variables. The table reports the parameter estimates along with the associated standard errors in parentheses. Average pricing errors and their standard errors are presented for each asset. The bottom row of each panel reports the test statistic for the overidentifying restrictions along with the associated finite-sample p-value in parentheses. For the identity weighting matrix, the "Dist" statistic has a non-standard aymptotic distribution with 90%, 95%, and 99% critical values given by 0.011, 0.016, and 0.025, respectively. For the efficient weighting matrix, the J-stat has an aymptotic chisquared distribution with 5 degrees of freedom. The 90%, 95%, and 99% critical values from the finite-sample distribution of the Dist statistic are 0.48, 0.59, and 0.95, respectively and those from that of the J-stat are 32.29, 41.78, and 54.64, respectively.

Table 4: Tests on the 6-Asset System

	Panel A: identity weighting matrix													
	μ	μ_d	ϕ	φ_d	ρ_1	ρ_0	φ_e	σ_1	σ_0	π_1	π_0	δ	γ	ψ
Estimates	0.015 (0.021)	0.014 (0.052)	3.0 (0.244)	4.0 (0.046)	0.92 (0.094)	0.8 (0.115)	0.5 (0.089)	0.011 (0.053)	0.025 (0.005)	0.7 (0.192)	0.9 (0.229)	0.98 (0.150)	10 (0.009)	0.6 (0.319)
Pricing Errors		*Mean*	*Std.Err.*											
Small		0.086	0.118											
Large		−0.007	0.088											
Growth		−0.007	0.090											
Value		0.064	0.109											
Market		0.001	0.089											
Risk free rate		−0.075	0.078											
$Dist-stat$	1.280 (>0.10)													
	Panel B: efficient weighting matrix													
	μ	μ_d	ϕ	φ_d	ρ_1	ρ_0	φ_e	σ_1	σ_0	π_1	π_0	δ	γ	ψ
Estimates	0.015 (0.003)	0.014 (0.008)	1.0 (0.364)	3.0 (0.274)	0.9 (0.354)	0.6 (0.297)	0.3 (0.171)	0.011 (0.066)	0.025 (0.010)	0.7 (0.040)	0.9 (0.019)	0.94 (0.106)	10 (0.370)	0.6 (0.341)
Pricing Errors		*Mean*	*Std.Err.*											
Small		0.107	0.101											
Large		0.002	0.057											
Growth		−0.003	0.057											
Value		0.089	0.098											
Market		0.012	0.061											
Risk free rate		−0.071	0.042											
$J-stat$	49.95 (>0.01)													

The table reports GMM estimates of the model using annual data over the period 1930-2006. Panels A and B report results for the identity and the efficient weighting matrices, respectively. The asset menu consists of the market portfolio, the risk free rate, and portfolios of "Small" capitalization, "Large" capitalization, "Growth", and "Value" stocks, giving 6 pricing restrictions. To this set of pricing restrictions, we add the 13 moment restrictions implied by the time-series specification of the model. This gives a total of 19 moment conditions. The total number of parameters to be estimated is 14, including 11 time-series and 3 preference parameters. The bigger root of the quadratic is used in the extraction of the latent state variables. The table reports the parameter estimates along with the associated standard errors in parentheses. Average pricing errors and their standard errors are presented for each asset. The bottom row of each panel reports the test statistic for the overidentifying restrictions along with the associated finite-sample p-value in parentheses. For the identity weighting matrix, the "Dist" statistic has a non-standard aymptotic distribution with 90%, 95%, and 99% critical values given by 0.035, 0.048, and 0.078, respectively. For the efficient weighting matrix, the J-statistic has an aymptotic chisquared distribution with 5 degrees of freedom. The 90%, 95%, and 99% critical values from the finite-sample distribution of the Dist statistic are 1.33, 1.70, and 2.05, respectively, and those from the finite-sample distribution of the J-statistic are 33.44, 40.05, and 53.72, respectively.

Table 5: Forecasting Regressions over 1930-2006

Panel A: Market Returns

const.	x	p	xp	xp^2	$log\,(P/D)$	R_f	R^2	*Adj*-R^2
−0.041	−2.384	0.165	77.59	−74.85			0.123	0.073
(0.053)	(1.277)	(0.060)	(76.53)	(76.64)				
0.417					−0.104		0.048	0.035
(0.178)					(0.054)			
−0.021					−0.112	0.461	0.065	0.039
(0.430)					(0.055)	(0.412)		

Panel B: Equity Premium

const.	x	p	xp	xp^2	$log\,(P/D)$	R_f	R^2	*Adj*-R^2
−1.019	−1.294	0.166	76.31	−73.61			0.151	0.103
(0.057)	(1.375)	(0.065)	(82.38)	(82.50)				
−0.659					−0.080		0.023	0.010
(0.197)					(0.060)			
−1.683					−0.099	1.076	0.098	0.073
(0.463)					(0.059)	(0.443)		

The table reports results from predictive regressions over the period 1930-2006. Panel A reports results for the market portfolio while Panel B does the same for the equity premium. The first row of Panel A (B) reports the regression coefficients along with the associated standard errors, the R^2, and the adjusted-R^2 from the forecasting regression of realized stock market returns (equity premium) on the two sate variables, x and p, their product, xp, and xp^2. The second row of Panel A (B) reports the regression coefficients along with the associated standard errors and the adjusted-R^2 from the forecasting regression of realized stock market returns (equity premium) on the lagged aggregate log price-dividend ratio. The third row of Panel A (B) reports the regression coefficients along with the associated standard errors and the adjusted-R^2 from the forecasting regression of realized stock market returns (equity premium) on the lagged aggregate log price-dividend ratio and the gross risk free rate.

Table 6: Forecasting Regressions over 1946-2006

Panel A: Market Returns

const.	x	p	xp	xp^2	$log\,(P/D)$	R_f	R^2	*Adj*-R^2
−0.023 (0.051)	−0.994 (1.569)	0.159 (0.059)	82.31 (67.28)	−81.81 (67.39)			0.207	0.150
0.471 (0.188)					−0.118 (0.056)		0.071	0.055
−0.488 (0.630)					−0.133 (0.056)	0.999 (0.627)	0.111	0.080

Panel B: Equity Premium

const.	x	p	xp	xp^2	$log\,(P/D)$	R_f	R^2	*Adj*-R^2
−0.998 (0.053)	0.217 (1.631)	0.161 (0.061)	83.54 (69.94)	−83.65 (70.06)			0.250	0.196
−0.533 (0.202)					−0.113 (0.060)		0.057	0.041
−2.098 (0.658)					−0.138 (0.058)	1.630 (0.655)	0.149	0.120

The table reports results from predictive regressions over the period 1946-2006. Panel A reports results for the market portfolio while Panel B does the same for the equity premium. The first row of Panel A (B) reports the regression coefficients along with the associated standard errors, the R^2, and the adjusted-R^2 from the forecasting regression of realized stock market returns (equity premium) on the two sate variables, x and p, their product, xp, and xp^2. The second row of Panel A (B) reports the regression coefficients along with the associated standard errors and the adjusted-R^2 from the forecasting regression of realized stock market returns (equity premium) on the lagged aggregate log price-dividend ratio. The third row of Panel A (B) reports the regression coefficients along with the associated standard errors and the adjusted-R^2 from the forecasting regression of realized stock market returns (equity premium) on the lagged aggregate log price-dividend ratio and the gross risk free rate.

Table 7: Out-of-Sample Forecasting

Out-of-Sample-R^2	
Market Portfolio	*Equity Premium*
0.058	0.109
0.005	0.022
0.014	0.069

The table reports results from out-of-sample predictive regressions over the period 1967-2006. Panel A reports results for the market portfolio while Panel B does the same for the equity premium. The first row of Panel A (B) reports the the out-of-sample-R^2 from the forecasting regression of realized stock market returns (equity premium) on the two sate variables, x and p, their product, xp, and xp^2. The second row of Panel A (B) reports the out–of-sample-R^2 from the forecasting regression of realized stock market returns (equity premium) on the lagged aggregate log price-dividend ratio. The third row of Panel A (B) reports the out–of-sample-R^2 from the forecasting regression of realized stock market returns (equity premium) on the lagged aggregate log price-dividend ratio and the gross risk free rate.

Table 8: Forecasting Regressions over 1930-2006 for Size-Sorted Portfolios

Panel A: Small Portfolio

const.	x	p	xp	xp^2	$log\,(P/D)$	R_f	R^2	*Adj*-R^2
−0.045 (0.095)	−5.293 (2.270)	0.203 (0.107)	168.7 (136.0)	−162.8 (136.2)			0.101	0.049
0.656 (0.314)					−0.164 (0.096)		0.039	0.025
0.670 (0.766)					−0.164 (0.097)	−0.015 (0.733)	0.039	0.012

Panel B: Large Portfolio

const.	x	p	xp	xp^2	$log\,(P/D)$	R_f	R^2	*Adj*-R^2
−0.043 (0.051)	−2.186 (1.231)	0.161 (0.058)	73.64 (73.76)	−70.94 (73.87)			0.128	0.078
0.378 (0.173)					−0.095 (0.053)		0.042	0.029
−0.105 (0.416)					−0.103 (0.053)	0.508 (0.398)	0.064	0.038

The table reports results from predictive regressions over the period 1930-2006. Panel A reports results for the "Small" portfolio while Panel B does the same for the "Large" portfolio. The first row of Panel A (B) reports the regression coefficients along with the associated standard errors, the R^2, and the adjusted-R^2 from the forecasting regression of realized stock market returns (equity premium) on the two sate variables, x and p, their product, xp, and xp^2. The second row of Panel A (B) reports the regression coefficients along with the associated standard errors and the adjusted-R^2 from the forecasting regression of realized stock market returns (equity premium) on the lagged aggregate log price-dividend ratio. The third row of Panel A (B) reports the regression coefficients along with the associated standard errors and the adjusted-R^2 from the forecasting regression of realized stock market returns (equity premium) on the lagged aggregate log price-dividend ratio and the gross risk free rate.

Table 9: Forecasting Regressions over 1930-2006 for B/M-Sorted Portfolios

Panel A: Growth Portfolio

const.	x	p	xp	xp^2	$log\,(P/D)$	R_f	R^2	$Adj\text{-}R^2$
-0.081 (0.057)	-2.762 (1.378)	0.198 (0.065)	110.6 (82.61)	-108.3 (82.73)			0.153	0.105
0.386 (0.197)					-0.100 (0.060)		0.036	0.023
0.012 (0.477)					-0.106 (0.061)	0.393 (0.457)	0.046	0.020

Panel B: Value Portfolio

const.	x	p	xp	xp^2	$log\,(P/D)$	R_f	R^2	$Adj\text{-}R^2$
0.003 (0.084)	-3.037 (2.018)	0.153 (0.095)	95.28 (121.0)	-91.39 (121.1)			0.059	0.005
0.585 (0.273)					-0.145 (0.083)		0.040	0.027
0.050 (0.661)					-0.154 (0.084)	0.563 (0.633)	0.050	0.024

The table reports results from predictive regressions over the period 1930-2006. Panel A reports results for the "Growth" portfolio while Panel B does the same for the "Value" portfolio. The first row of Panel A (B) reports the regression coefficients along with the associated standard errors, the R^2, and the adjusted-R^2 from the forecasting regression of realized stock market returns (equity premium) on the two sate variables, x and p, their product, xp, and xp^2. The second row of Panel A (B) reports the regression coefficients along with the associated standard errors and the adjusted-R^2 from the forecasting regression of realized stock market returns (equity premium) on the lagged aggregate log price-dividend ratio. The third row of Panel A (B) reports the regression coefficients along with the associated standard errors and the adjusted-R^2 from the forecasting regression of realized stock market returns (equity premium) on the lagged aggregate log price-dividend ratio and the gross risk free rate.

Table 10: Forecasting Regressions over 1946-2006 for Size-Sorted Portfolios

Panel A: Small Portfolio

const.	x	p	xp	xp^2	$log\,(P/D)$	R_f	R^2	$Adj\text{-}R^2$
0.038	−0.767	0.100	172.3	−174.7			0.111	0.046
(0.085)	(2.608)	(0.098)	(111.8)	(112.0)				
0.348					−0.075		0.012	−0.005
(0.304)					(0.090)			
−0.190					−0.084	0.560	0.017	−0.018
(1.039)					(0.092)	(1.034)		

Panel B: Large Portfolio

const.	x	p	xp	xp^2	$log\,(P/D)$	R_f	R^2	$Adj\text{-}R^2$
−0.029	−1.042	0.167	79.00	−77.80			0.212	0.154
(0.051)	(1.558)	(0.058)	(66.78)	(66.89)				
0.469					−0.118		0.073	0.057
(0.186)					(0.055)			
−0.648					−0.137	1.163	0.127	0.096
(0.621)					(0.055)	(0.618)		

The table reports results from predictive regressions over the period 1946-2006. Panel A reports results for the "Small" portfolio while Panel B does the same for the "Large" portfolio. The first row of Panel A (B) reports the regression coefficients along with the associated standard errors, the R^2, and the adjusted-R^2 from the forecasting regression of realized stock market returns (equity premium) on the two sate variables, x and p, their product, xp, and xp^2. The second row of Panel A (B) reports the regression coefficients along with the associated standard errors and the adjusted-R^2 from the forecasting regression of realized stock market returns (equity premium) on the lagged aggregate log price-dividend ratio. The third row of Panel A (B) reports the regression coefficients along with the associated standard errors and the adjusted-R^2 from the forecasting regression of realized stock market returns (equity premium) on the lagged aggregate log price-dividend ratio and the gross risk free rate.

Table 11: Forecasting Regressions over 1946-2006 for B/M-Sorted Portfolios

Panel A: Growth Portfolio

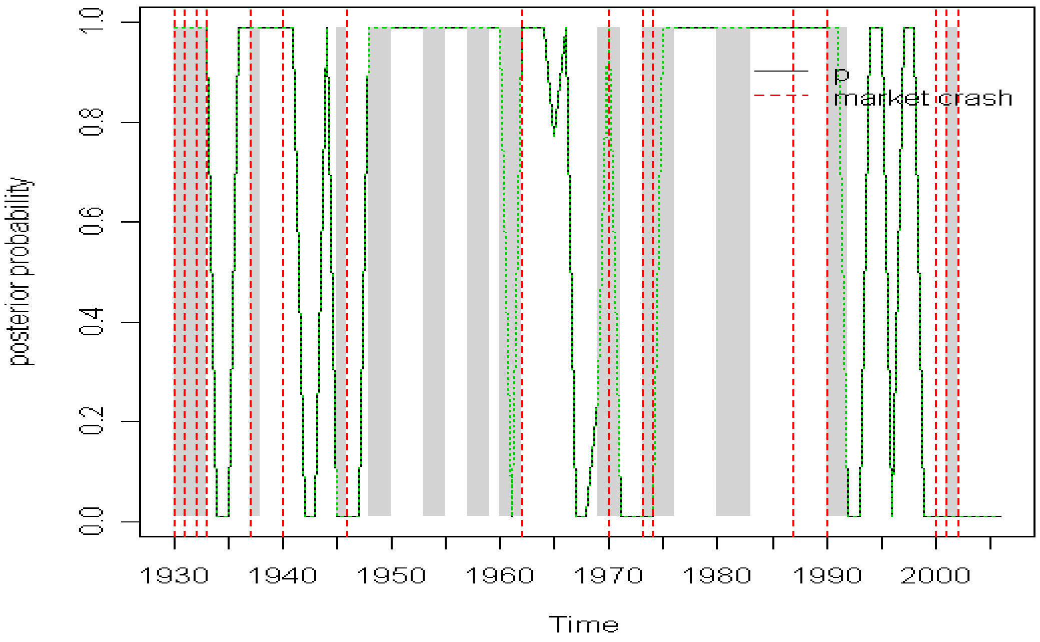
const.	x	p	xp	xp^2	$log\,(P/D)$	R_f	R^2	*Adj*-R^2
−0.072 (0.059)	−1.761 (1.806)	0.204 (0.068)	114.5 (77.43)	−114.0 (77.56)			0.243	0.188
0.486 (0.222)					−0.127 (0.066)		0.059	0.043
−0.391 (0.753)					−0.141 (0.067)	0.914 (0.749)	0.083	0.051

Panel B: Value Portfolio

const.	x	p	xp	xp^2	$log\,(P/D)$	R_f	R^2	*Adj*-R^2
0.051 (0.073)	−0.397 (2.256)	0.115 (0.085)	104.9 (96.73)	−105.6 (96.90)			0.098	0.032
0.497 (0.257)					−0.114 (0.077)		0.037	0.020
−0.469 (0.873)					−0.130 (0.078)	1.007 (0.869)	0.059	0.026

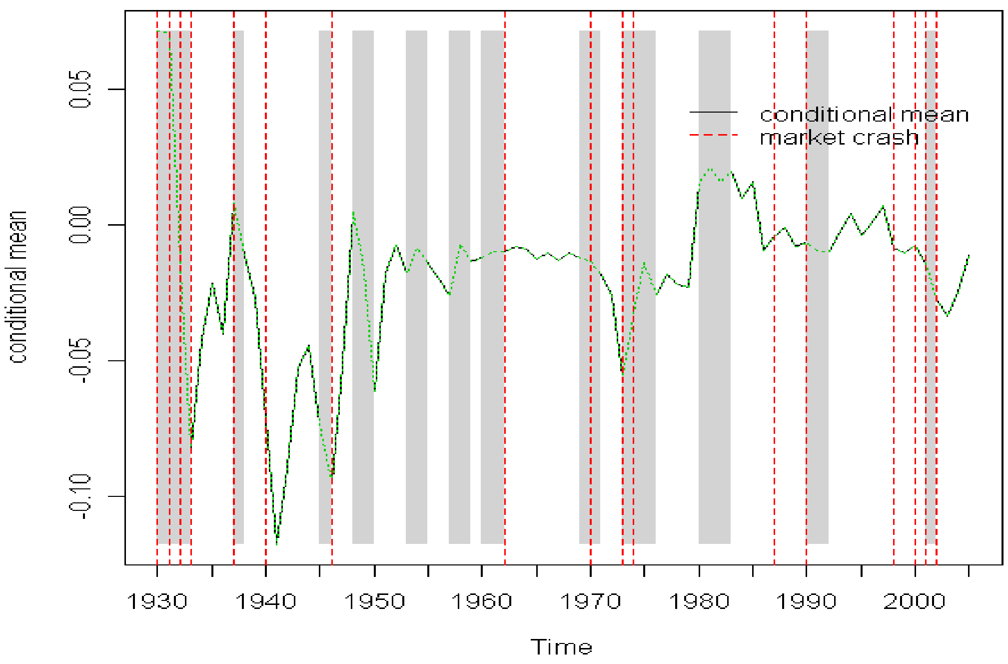
The table reports results from predictive regressions over the period 1946-2006. Panel A reports results for the "Growth" portfolio while Panel B does the same for the "Value" portfolio. The first row of Panel A (B) reports the regression coefficients along with the associated standard errors, the R^2, and the adjusted-R^2 from the forecasting regression of realized stock market returns (equity premium) on the two sate variables, x and p, their product, xp, and xp^2. The second row of Panel A (B) reports the regression coefficients along with the associated standard errors and the adjusted-R^2 from the forecasting regression of realized stock market returns (equity premium) on the lagged aggregate log price-dividend ratio. The third row of Panel A (B) reports the regression coefficients along with the associated standard errors and the adjusted-R^2 from the forecasting regression of realized stock market returns (equity premium) on the lagged aggregate log price-dividend ratio and the gross risk free rate.

Figure 1: Posterior Probability of First Regime



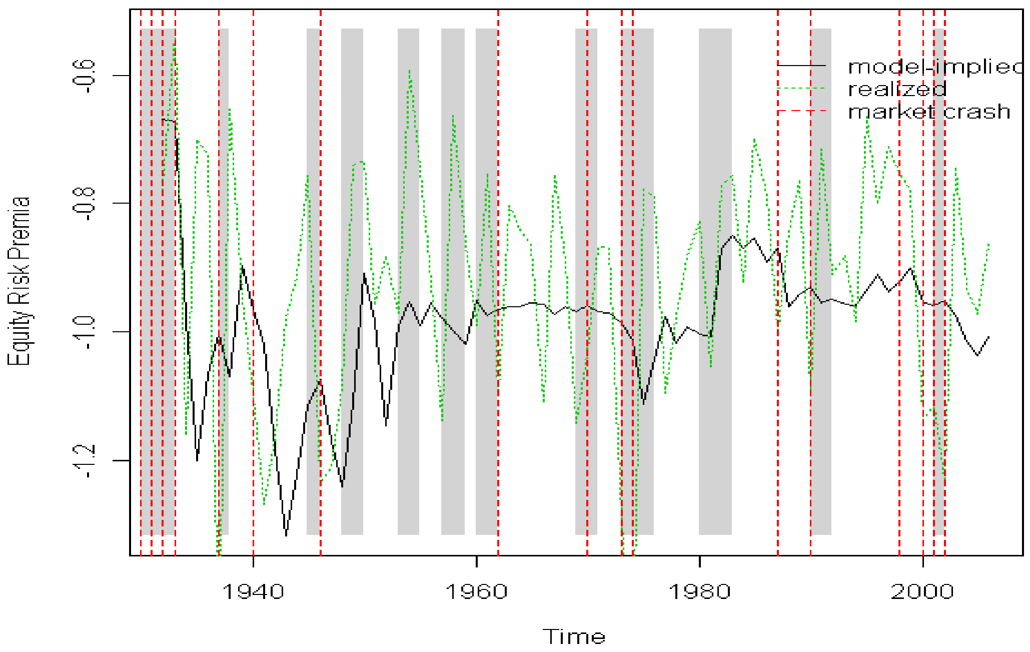
The Figure plots the extracted time-series of the consumer's posterior probability of being in the first regime over 1930-2006. The aggregate log price-dividend ratio and the risk free rate are affine functions only of the two state variables and their product. Hence, we can invert the system to extract the two state variables as known functions of the observable aggregate log price-dividend ratio and interest rate. This procedure of inverting two non-linear equations yields quadratic equations for x_t and p_t, with coefficients that depend on $z_{m,t}$ and $R_{f,t}$, and the time-series and preference parameters. Solving the equations gives two pairs of solutions for x_t and p_t as functions of the observables, $z_{m,t}$ and $R_{f,t}$. We use the bigger root of the quadratic equations to extract the latent state variables as this choice minimizes the value of the criterion function. The SME point estimates of the time-series and preference parameters are used in the extraction of the latent state variables.

Figure 2: Conditional Mean of Consumption Growth



The Figure plots the extracted time-series of the conditional mean of aggregate consumption and dividend growth over 1930-2006. The aggregate log price-dividend ratio and the risk free rate are affine functions only of the two state variables and their product. Hence, we can invert the system to extract the two state variables as known functions of the observable aggregate log price-dividend ratio and interest rate. This procedure of inverting two non-linear equations yields quadratic equations for x_t and p_t, with coefficients that depend on $z_{m,t}$ and $R_{f,t}$, and the time-series and preference parameters. Solving the equations gives two pairs of solutions for x_t and p_t as functions of the observables, $z_{m,t}$ and $R_{f,t}$. We use the bigger root of the quadratic equations to extract the latent state variables as this choice minimizes the value of the criterion function. The SME point estimates of the time-series and preference parameters are used in the extraction of the latent state variables.

Figure 3: Model-Implied and Realized Equity Risk Premia



The Figure plots the realized time-series of the equity premium along with the model-implied time-series of the same over 1930-2006. Also shown in the graph are the NBER-registered recessions (shaded areas) and the major stock market crashes (dashed lines).