The Predictability of Returns with Regime Shifts in Consumption and Dividend Growth*

Anisha Ghosh† George M. Constantinides‡

this version: May 2, 2011

Abstract

We present evidence that the stock market return, dividend growth, and consumption growth are predictable. The key insight is that the consumption and dividend growth processes differ across two latent economic regimes. We estimate the equilibrium model and identify the probability that the economy is in the first regime as a non-linear function of the risk free rate and market-wide price-dividend ratio. The second regime is associated with recessions, market downturns, and higher volatility of returns and growth rates. The model-implied state variables perform significantly better at in-sample forecasting and out-of-sample prediction of the equity, size, and value premia and consumption and dividend growth rates than the price-dividend ratio and risk free rate.

Keywords: Return Predictability, Consumption Growth Predictability, Dividend Growth Predictability, Regime Shifts, Cross-Section of Returns, Equity Premium, Size Premium, Value Premium.

JEL classification: G12, E44

*We thank Raj Chakrabarti, John Cochrane, Rick Green, Lars Hansen, Burton Hollifield, Christian Julliard, Ralph Koijen, Oliver Linton, Monika Piazzesi, and Bryan Routledge and seminar participants at Carnegie Mellon University, the University of Chicago, Stanford GSB, the Sixth Annual Early Career Women in Finance Mini-Conference, and the Stanford Institute for Theoretical Economics 2010 Workshop for helpful comments. We remain responsible for errors and omissions. Constantinides acknowledges financial support from the Center for Research in Security Prices of the University of Chicago Booth School of Business.

†Tepper School of Business, Carnegie Mellon University; anishagh@andrew.cmu.edu, https://wpweb2.tepper.cmu.edu/rlang/faculty/AnishaGhoshWeb/default.html.

‡Corresponding author; Booth School of Business, University of Chicago and NBER; 5807 South Woodlawn Avenue, Chicago IL 60637; Ph: 773-702-7258; Fax: 773-753-8045; gmc@chicagobooth.edu, http://faculty.chicagobooth.edu/george.constantinides/.

1 Introduction

The predictability of the aggregate stock market return, dividend growth, and consumption growth have for long been the subject of both theoretical and empirical research in economics and finance. In this paper, we present evidence that the aggregate stock market return, dividend growth, and consumption growth are predictable both in sample and out of sample. The key insight is that the aggregate consumption and dividend growth processes differ across (at least) two latent economic regimes. We estimate the equilibrium model and identify the probability that the economy is in one of two economic regimes as a non linear function of two financial variables, the short term risk free rate and the market-wide price-dividend ratio. The regimes are related to the business cycle: the probability of a recession in a year is 44.4% (8.2%), if the probability of being in the second (first) regime at the beginning of the year exceeds 50%. The regimes are also related to the major stock market downturns, as identified in Barro and Ursua (2009): the probability of a stock market downturn in a year is 44.4% (8.2%), if the probability of being in the second (first) regime at the beginning of the year exceeds 50%.

Attempts to predict the aggregate stock market return have a long history in economics going back to as early as 1920 when Dow (1920) explored the role of dividend ratios in predicting the market return. Over the last three decades, the academic literature has explored numerous macroeconomic and financial variables as potential predictors of the market return and equity premium. The price-dividend ratio has received extensive scrutiny as a predictive variable because, as a mathematical identity, all variation in the price-dividend ratio must be accounted for by changing expectations on future returns and/or future dividend growth (Campbell and Shiller (1988)).[1] Welch and Goyal (2008) review this literature and undertake a comprehensive study of the in-sample and out-of-sample performance of the price-dividend ratio and other variables in predicting the equity premium. They conclude that "by and large these models have predicted poorly both in sample and out of sample for 30 years now; these models seem unstable, as diagnosed by their out-of-sample predictions and other statistics; and these models would not have helped an investor with access only to available information to profitably time the market." These conclusions are controversial.

The benchmark in the consumption growth literature is Hall's (1978) demonstration that, under rational expectations and time- and state-separable preferences, marginal utility of consumption is unpredictable. The statistical difficulty in distinguishing between an i.i.d. process from one with a small but persistent predictable component has led to considerable controversy regarding the presence of a predictable component in the aggregate consumption growth rate. The debate on consumption growth predictability remains open.

In this paper, we shed light on this debate by arguing that there exist (at least) two latent economic regimes. The predictable component of the aggregate consumption and

[1] See also Ang and Bekaert (2007), van Binsbergen and Koijen (2010), Boudoukh Richardson, and Whitelaw (2008), Campbell and Thompson (2008), Cochrane (2008), Fama and French (1988), Kelly and Pruitt (2010), and Lettau and Van Nieuwerburgh (2008)).

dividend growth rates has persistence and volatility that is different across the regimes. We identify the regimes in the context of a dynamic equilibrium asset pricing model with two regimes. The probability that the economy is in the first regime is obtained as a non linear function of the market-wide price-dividend ratio and short term risk free rate, with parameters estimated from the Euler equations of the market return, risk free rate, and the cross-section of size- and book-to-market-equity-sorted portfolio returns, plus unconditional moments of the consumption and dividend processes. Furthermore, this non-linearity cannot be captured by simple nonlinear functions like a quadratic function of the market-wide price-dividend ratio and risk free rate.

Over the period $1930 - 2009$, in all years when the probability of being in the second regime exceeds 50%, an in-sample linear forecasting regression of the realized aggregate consumption growth rate on the lagged market-wide price-dividend ratio yields a statistically significant coefficient and (adjusted) $\overline{R}^2$ 41.3%; the regression on the aggregate dividend growth yields a statistically significant coefficient and (adjusted) $\overline{R}^2$ 23.7%. The price-dividend ratio performs poorly at predicting the market return and equity premium in the second regime with $\overline{R}^2$ -5.9% and -1.3%, respectively. The converse is true in the first regime. The forecasting power of the price-dividend ratio for the consumption and dividend growth rates is poor with $\overline{R}^2$ -1.0% and -1.6%, respectively. The regressions of the realized one-year real market return and equity premium on the price-dividend ratio have coefficients of the right sign and (adjusted) $\overline{R}^2$ 1.4% and 1.5%, respectively.

In the model, a state variable x_t that simultaneously drives the conditional means of the aggregate consumption and dividend growth rates reverts to its unconditional mean with a process that differs across two regimes. Based on his information set, the consumer observes x_t and also calculates the probability, p_t, that the economy is in the first regime. The conditional means of the aggregate consumption and dividend growth rates are affine functions of the two state variables $(x_t,\, p_t)$. The market-wide log price-dividend ratio and risk free rate are approximately affine functions of $(x_t,\, p_t)$ and their product, thereby rendering the (potentially latent) state variables and the expected return of each asset class known nonlinear functions of the price-dividend ratio and risk free rate. The model parameters are estimated from the Euler equations of the market return, risk free rate, and the cross-section of size and book-to-market-equity sorted portfolio returns plus unconditional moments of the consumption and dividend processes.

We show that the model has superior in-sample forecasting performance for the equity premium and its variance relative to a linear forecasting model with the market-wide price-dividend ratio and risk free rate as predictive variables. Moreover, unlike linear forecasting regressions with the price-dividend ratio and risk free rate as predictive variables, the model-implied state variables have robust forecasting performance across subperiods.

While most of the predictability literature focuses on predicting the aggregate US stock market return and equity premium, the literature on the time series forecastability of the cross-section of size and book-to-market-equity sorted portfolio returns is scant.[2] Forecastability of the cross-section of returns is important for at least two reasons. First, the historical size premium (9.4%) and value premium (7.3%) are of the same order of magnitude as the equity premium (7.9%), based on arithmetic annual returns. Therefore, the predictability of these premia is important in active portfolio management. Second, it is also important in providing an alternative channel to examine the empirical plausibility of a given set of state variables that purport to explain the cross-section of returns. We show that the model has superior forecasting performance for the size and value premia relative to the linear forecasting model; furthermore this performance is robust across subperiods.

We demonstrate that our model retains its predictive power out of sample. The model-implied state variables give an out-of-sample R^2 of 5.2%, 22.6%, and 0.0%, respectively, for the equity, size, and value premia over the period $1976 - 2009$. When used as predictive variables in a linear predictive model, the price-dividend ratio and risk free rate have poor predictive performance with out-of-sample R^2 of -2.6%, -6.8%, and -11.9%, respectively, for the equity, size, and value premia.

Finally, we show that the model-implied state variables have strong forecasting power for the aggregate consumption and dividend growth rates and their variances. In-sample forecasting regressions for the consumption and dividend growth rates give statistically significant coefficients on the state variables and $\overline{R}^2$ 8.0% and 11.7%, respectively. The corresponding $\overline{R}^2$ are 12.7% and 15.5%, respectively, for the conditional variances of the growth rates. The model-implied state variables give an out-of-sample R^2 of 0.7% and 4.4%, respectively, for the consumption and dividend growth rates over the period $1976 - 2009$. The price-dividend ratio and risk free rate perform poorly at predicting the consumption and dividend growth rates out of sample, giving large negative R^2 of -123.2% and -60.3%, respectively. These results provide strong support for the risk channels highlighted in the model and the precise mechanism by which they drive the dynamics of the consumption and dividend growth processes.

Our paper is related to equilibrium models by Bansal and Shaliastovich (2011), Bansal and Yaron (2004), Drechsler (2009), Hansen, Heaton and Li (2008), Hore (2010), Lettau and Ludvigson (2001), and Menzly, Santos, and Veronesi (2004) with implications on forecasting the market return and dividend growth. Our paper is also related to Brandt and Kang (2004), van Binsbergen and Koijen (2010), Kelly and Pruitt (2010), and Pastor and Stambaugh (2009) who focus on return predictability using filtering techniques. While these are reduced form models, we rely on an equilibrium model and avoid using filtering techniques by arguing that, under the model assumptions, the (potentially latent) state variables and the expected return of each asset class are known nonlinear functions of observable financial variables like the price-dividend ratio and risk free rate. Constantinides and Ghosh (2011) earlier applied a similar inversion methodology to extract latent state variables in the context of the Bansal and Yaron (2004) long run risks model and its cointegrated extension.

Our paper is also related to Lettau and Van Nieuwerburgh (2008), Pastor and Stambaugh (2001), and Paye and Timmermann (2006) who find evidence of structural

[2] There are two notable exceptions. First, Baker and Wurgler (2006) find that the cross-section of future stock returns is conditional on beginning-of-period proxies for sentiment. Second, the January dummy has strong predictive power for size and book-to-market-equity sorted portfolio returns.

breaks and argue that allowance for these breaks has important implications for return predictability. Finally, our paper is related to Constantinides, Jackwerth, and Savov (2011) who highlight the importance of regime shifts by finding that a pricing factor that tracks jumps in the volatility of the market return explains the cross-section of index option returns; it also explains the cross-section of equity returns as well as the SMB factor and almost as well as the HML factor.

The paper is organized as follows. In Section 2, we present the regime shifts model. We express the price-dividend ratio, risk free rate, expected equity premium, and expected consumption and dividend growth rates as functions of the state variables $(x_t,\ p_t)$. In Section 3, we discuss the data. In Section 4, we estimate the model parameters with GMM from the set of the Euler equations for the market return, risk free rate, and portfolios of "Small", "Large", "Growth" and "Value" stocks, and the unconditional moments of the consumption and dividend growth processes. Using the point estimates of the model parameters, we invert the expressions for the price-dividend ratio and risk free rate as functions of the state variables and express the state variables as functions of the price-dividend ratio and risk free rate.

Armed with the time series of the state variables, we address the questions raised in this paper. Section 5 presents empirical evidence that the predictability of returns and the aggregate consumption and dividend growth rates differ significantly in the two-regimes. In Section 6, we present evidence on the in-sample and out-of-sample predictability of the equity, size, and value premia. In Section 7, we present evidence on the in-sample and out-of-sample predictability of the aggregate consumption and dividend growth rates. In Section 8, we present evidence on the predictability of the variance of the market return and the growth rate of consumption and dividends. Section 9 concludes. The Appendix contains the derivation of the main results.

2 The Model and Implications for Predictability

We present the regime shifts model and its implications for the predictability of the equity, size, and value premia and consumption and dividend growth.

2.1 Model

The model stipulates that the state variable, x_t, that simultaneously drives the conditional means of the aggregate consumption and dividend growth rates reverts to its unconditional mean with a process that differs across two regimes:

$$x_{t+1} = \rho_{s_{t+1}} x_t + \varphi_e \sigma_{s_{t+1}} e_{t+1}, \tag{1}$$
$$\Delta c_{t+1} = \mu + x_t + \sigma_{s_{t+1}} \eta_{t+1}, \tag{2}$$
$$\Delta d_{t+1} = \mu_d + \phi x_t + \varphi_d \sigma_{s_{t+1}} u_{t+1}, \tag{3}$$

where c_{t+1} is the logarithm of the aggregate consumption level; d_{t+1} is the logarithm of the aggregate stock market dividends; and $s_t = 1,\ 2$ is a variable that denotes the economic regime. The persistence parameter, ρ_{s_t}, of the state variable x_t and the level of its volatility, σ_{s_t}, are generally different in the two regimes. The shocks e_{t+1}, η_{t+1}, and u_{t+1} are assumed to be distributed with mean 0 and variance 1 and independent of the past.

Given his information set, $\mathcal{F}(t)$, the representative consumer observes x_t and calculates his probability, p_t, at time t of being in regime $s_t = 1$:

$$p_t \equiv Prob\left(s_t = 1 | \mathcal{F}(t)\right) \tag{4}$$

We do not take a stand on the content of the information set, $\mathcal{F}(t)$. In one extreme case, it may be limited to the history of consumption, dividends, and past realizations of x. In the other extreme case, it may include all publicly available information. Furthermore, we do not take a stand on the optimality of the filter that the consumer applies to form his belief, p_t. The econometrician does not directly observe the state variables, p_t and x_t, and, hence, they are latent.

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

$$\Pi = \begin{pmatrix} \pi_1 & 1-\pi_2 \\ 1-\pi_1 & \pi_2 \end{pmatrix}, \tag{5}$$

where $0 < \pi_i < 1$ for $i = 1, 2$. Thus, the consumer's probability of being in regime $s_{t+1} = 1$ at time $t+1$, given his information set, $\mathcal{F}(t)$, is

$$Prob\left(s_{t+1} = 1 | \mathcal{F}(t)\right) = \pi_1 p_t + (1-\pi_2)(1-p_t) \equiv f(p_t). \tag{6}$$

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

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

$$p_{t+1} = f(p_t) + \varepsilon_{t+1}, \tag{7}$$

where $E\left[\varepsilon_{t+1} | \mathcal{F}(t)\right] = 0$.

We make the following assumptions regarding the shocks η_{t+1}, u_{t+1}, e_{t+1}, and ε_{t+1}:

$$E\left[y_{t+1} | \mathcal{F}(t), s_{t+1} = 1\right] = E\left[y_{t+1} | s_{t+1} = 1\right] \equiv y(1), \text{ a constant,} \tag{8}$$

where $y = \eta$, u, e, and ε;

$$E\left[y_{t+1} w_{t+1} | \mathcal{F}(t), s_{t+1} = 1\right] = E\left[y_{t+1} w_{t+1} | \mathcal{F}(t)\right] \equiv \sigma_{y,w}, \text{ a constant,} \tag{9}$$

where y, $w = \eta$, u, e, ε, and $y \neq w$; and

$$E\left[y_{t+1}^2 | \mathcal{F}(t), s_{t+1} = 1\right] = E\left[y_{t+1}^2\right] = 1, \tag{10}$$

where $y = \eta$, u, and e.

Equation (8) recognizes that the means of the residuals η_{t+1}, u_{t+1}, e_{t+1}, and ε_{t+1}, conditional on the regime at time $t+1$, may differ from their unconditional value of zero. To ensure that p_t lies in the permissible interval $[0,1]$, we restrict $\varepsilon(1)$ in equation (8) such that

$$\varepsilon(1) \in \begin{cases} \left[\max\left(-(1-\pi_2), -\frac{(1-\pi_1)^2}{1-\pi_2}\right), 1-\pi_1\right] & , \text{ if } \pi_1+\pi_2-1>0, \\ \left[\max\left(-\pi_1, -\frac{(1-\pi_1)\pi_2}{1-\pi_2}\right), \min\left(\pi_2, \frac{(1-\pi_1)\pi_1}{1-\pi_2}\right)\right], & \text{ if } \pi_1+\pi_2-1<0, \end{cases} \tag{11}$$

(see Appendix A.1 for derivation of this result). Equation (9) recognizes that the residuals η_{t+1}, u_{t+1}, e_{t+1}, and ε_{t+1} may be correlated. Finally, equation (10) limits the number of parameters to be estimated by setting the second moments of the residuals η_{t+1}, u_{t+1}, and e_{t+1}, conditional on the regime at time $t+1$, equal to their unconditional value of one.

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

$$V_t = \left[(1-\delta)C_t^{\frac{1-\gamma}{\theta}} + \delta\left(E\left[V_{t+1}^{1-\gamma}|\digamma(t)\right]\right)^{\frac{1}{\theta}}\right]^{\frac{\theta}{1-\gamma}}, \tag{12}$$

where δ denotes the subjective discount factor, $\gamma > 0$ is the coefficient of risk aversion, $\theta \equiv \frac{1-\gamma}{1-\frac{1}{\psi}}$, 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,

$$E\left[\exp(m_{t+1} + r_{j,t+1})|\digamma(t)\right] = 1, \tag{13}$$

$$m_{t+1} = \theta\log\delta - \frac{\theta}{\psi}\Delta c_{t+1} + (\theta-1)r_{c,t+1}, \tag{14}$$

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 return on the consumption claim, $r_{c,t+1}$, and that on the market portfolio (the observable return on the aggregate dividend claim), $r_{m,t+1}$, as in Campbell and Shiller (1988),

$$\begin{aligned} r_{c,t+1} &= \kappa_0 + \kappa_1 z_{t+1} - z_t + \Delta c_{t+1}, && (15) \\ r_{m,t+1} &= \kappa_{0,m} + \kappa_{1,m} z_{m,t+1} - z_{m,t} + \Delta d_{t+1}, && (16) \end{aligned}$$

where z_t is the log price-consumption ratio and $z_{m,t}$ the log price-dividend ratio. In equation (15), $\kappa_1 = \frac{e^{\overline{z}}}{1+e^{\overline{z}}}$ 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 (16), $\kappa_{1,m} = \frac{e^{\overline{z}_m}}{1+e^{\overline{z}_m}}$ and $\kappa_{0,m} = \log(1 + e^{\overline{z}_m}) - \kappa_{1,m}\overline{z}_m$, 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 the following approximate expressions for the log price-consumption ratio, log price-dividend ratio and log risk free rate and derive expressions for their parameters in Appendices A.2.1, A.2.2, and A.2.3, respectively:

$$
\begin{aligned}
z_t &= p_t \left[A_0(1) + A_1(1)x_t\right] + (1 - p_t)\left[A_0(2) + A_1(2)x_t\right], && (17) \\
z_{m,t} &= p_t \left[A_{0,m}(1) + A_{1,m}(1)x_t\right] + (1 - p_t)\left[A_{0,m}(2) + A_{1,m}(2)x_t\right], && (18) \\
r_{f,t} &= A_{0,f} + A_{1,f}x_t + A_{2,f}p_t + A_{3,f}p_t x_t. && (19)
\end{aligned}
$$

The pricing kernel is a function of the two latent state variables, x_t and p_t, and their lags, in addition to consumption growth (see Appendix A.2.6 for derivation). We invert the two non-linear equations (18) and (19) to express the latent state variables, x_t and p_t, as functions of the observables, $z_{m,t}$ and $r_{f,t}$. This gives a quadratic equation for p_t, with coefficients that depend on $z_{m,t}$ and $r_{f,t}$, and the time-series and preference parameters. We set p_t equal to the bigger root of the quadratic equation.[3] Finally, we obtain x_t which is given as a function of p_t. This procedure gives a pricing kernel entirely in terms of observables. We use this expression for the pricing kernel to estimate the parameters of the model using a cross-section of asset returns in Section 4. Figures 1 and 2 display p and x, respectively, as highly non-linear functions of z_m and r_f using the point estimates of the model parameters in Section 4.[4]

— Figures 1 and 2 about here —

[3] We justify this choice via simulation. Specifically, we calibrate the model using the point estimates of the parameters in Section 4, generate a time series of x_t and p_t of length 10,000 years, and then generate the time series of $z_{m,t}$ and $r_{f,t}$. Each year, we obtain the quadratic equation of p_t, with coefficients that depend on the generated values of $z_{m,t}$ and $r_{f,t}$. We invariably find that the known value of p_t that generated the time series equals the bigger root of the quadratic equation. In the data, due to parameter estimation error, we infrequently find that the bigger root is slightly greater than one and, in this case, set it equal to 0.99; we also infrequently find that the bigger root is slightly smaller than zero and set it equal to 0.01.

[4] At the point estimates of the parameters, we obtain the following quadratic equation for p_t at time t:

$$a p_t^2 + b_t p_t + h_t = 0,$$

where

$$
\begin{aligned}
a &= -2.3 \times 10^{-14}, \\
b_t &= 8.94 r_{f,t} + 0.24, \\
h_t &= 5.63 r_{f,t} - 1.11 z_{m,t} + 2.67,
\end{aligned}
$$

and

$$x_t = \frac{r_{f,t} - 0.001 - \left(2.6 \times 10^{-15}\right) p_t}{1.11}.$$

2.2 Predictive Implications for Returns and Growth Rates

Equations (16), (18), and (3) imply that the expected market return is given by:

$$E\left[r_{m,t+1}|\digamma(t)\right] = B_0 + B_1 x_t + B_2 p_t + B_3 p_t x_t. \tag{20}$$

Hence, from Equations (20) and (19), the expected equity premium is given by:

$$\begin{aligned} E\left[\left(r_{m,t+1} - r_{f,t}\right)|\digamma(t)\right] &= E_0 + E_1 x_t + E_2 p_t + E_3 p_t x_t, \\ E_i &= B_i - A_{i,f}, \quad i = 0, 1, ..., 3. \end{aligned} \tag{21}$$

Therefore, the model generates time-varying expected market return and equity premium. The coefficients $\{B_i, E_i\}_{i=0}^{3}$ are known functions of the underlying time-series and preference parameters (see Appendix A.2.4 for derivation). Under the assumption that the dividend growth processes of the "Small", "Large", "Growth" and "Value" portfolios are similar to that for the market, the expected returns on these portfolios can also be shown to be affine functions of the state variables, x and p, and their product.

The regime shifts model also has implications for the predictability of the aggregate consumption and dividend growth rates (see Appendix A.2.5 for derivation). The time series specification of the model implies that the expected consumption growth rate is given by

$$E\left(\Delta c_{t+1}|\digamma(t)\right) = \mu + \left(\sigma_1 - \sigma_2\right)\eta(1)\left(1 - \pi_2\right) + x_t + \left(\sigma_1 - \sigma_2\right)\eta(1)\left(\pi_1 + \pi_2 - 1\right)p_t, \tag{22}$$

and the expected dividend growth rate is given by

$$E\left(\Delta d_{t+1}|\digamma(t)\right) = \mu_d + \varphi_d\left(\sigma_1 - \sigma_2\right)u(1)\left(1 - \pi_2\right) + \phi x_t + \varphi_d\left(\sigma_1 - \sigma_2\right)u(1)\left(\pi_1 + \pi_2 - 1\right)p_t, \tag{23}$$

both linear functions of the state variables, x_t and p_t.

Finally, the model implies that the conditional variance of the aggregate consumption and dividend growth rates are functions of the probability, p_t, alone:

$$Var\left(\Delta c_{t+1}|\digamma(t)\right) = a_c + b_c p_t, \tag{24}$$

and

$$Var\left(\Delta d_{t+1}|\digamma(t)\right) = a_d + b_d p_t. \tag{25}$$

3 Data

We consider the predictive performance of the model at the annual frequency, using annual data over the entire available sample period $1930 - 2009$. The asset menu consists of the market return, 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 annual real risk free rate is the inflation-adjusted rolled-over return of one-month Treasury Bills from Ibbotson Associates. The equity premium is the difference in average returns on the market and the risk free rate.

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. The size premium is the difference in average returns between the "Small" and "Large" portfolios. 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. The value premium is the difference in average returns between the "Value" and "Growth" portfolios. Annual returns for the "Small", "Large", "Growth", and "Value" portfolios are computed by compounding monthly returns within each year. The premia are computed as the difference in the average annual returns.

Also used in the empirical analysis are the price-dividend ratio and dividend growth rate of the market portfolio. These two time series are computed using the monthly returns with and without dividends on the market portfolio obtained from the CRSP files. The monthly dividend payments within a year are added to obtain the annual aggregate dividend, i.e. we do not reinvest dividends either in T-Bills or in the aggregate stock market. The annual price-dividend ratio is computed as the ratio of the price at the end of each calender year to the annual aggregate dividends paid out during that year.

Finally, the consumption data consists of the per capita personal consumption expenditure on nondurable goods obtained from the Bureau of Economin Analysis. All nominal quantities are converted to real, using an $ARMA(1,1)$ forecast of the annual inflation.

4 Parameter Estimation and Interpretation

We estimate the model parameters over the period $1930-2009$ with GMM on the following set of 23 moment conditions, weighted by the identity matrix: the 18 Euler equations for the risk free rate, market return, and "Small", "Large", "Growth", and "Value" portfolio returns, with the risk free rate and the lagged log price-dividend ratio of the market as instruments; and the five moment restrictions implied by the unconditional means and variances of the aggregate consumption and dividend growth rates and the covariance between consumption and dividend growth rates (see Appendix A.3 for derivation of these moments). The total number of parameters to be estimated is 21: 3 preference parameters (γ, ψ, δ); 16 time-series parameters (μ, μ_d, ϕ, φ_d, ρ_1, ρ_2, σ_1, σ_2, π_1, π_2, φ_e, $e(1)$, $\eta(1)$, $u(1)$, $\varepsilon(1)$, $\sigma_{\varepsilon,e}$); and 2 combinations of all the parameters that appear in the Euler equations.

The estimation results are reported in Table I. The first and second rows report the point estimates of the parameters along with the associated standard errors in parentheses.[5] The point estimates of the subjective discount factor (0.976) and risk aversion coefficient (12) are economically sensible. The point estimate of the IES is 0.9 and is slightly smaller than one.

— Table I about here —

The parameter estimates of the time-series processes illustrate the presence of (at least) two regimes, with more persistent and less volatile consumption and dividend growth rates in the first regime than in the second one. The persistence parameter of the state variable x is 0.94 in the first regime (half-life longer than 11 years) and 0.60 in the second one (half life of just over one year); and the volatility of x is 0.5% in the first regime and 3.5% in the second one. The point estimates of the transition probabilities imply that the first regime has a mean duration of 20 years while the second regime has a much shorter duration of just over 6 years.

The estimates of the time-series parameters in Table I are consistent with the time-series specification of the model. The model generates almost perfectly the first two sample moments of consumption growth. The unconditional mean and volatility of the aggregate consumption growth rate are 1.5% and 2.5%, respectively, in the data. The median values for these moments obtained from 10000 simulated samples of the same length as the historical data are 1.5% and 2.4%, respectively (see Appendix A.4 for details of the simulation design). The model also does a good job in generating the sample correlation of consumption and dividend growth: the sample value of this correlation is 0.59 while the median value obtained through simulation is 0.67. The sample mean of dividend growth lies within the 95% confidence interval of the simulated moment. The sample standard deviation of dividend growth lies slightly above the 95% confidence interval of the simulated moment.

The model also does a good job of matching the means of the risk free rate, equity premium, and the market-wide price-dividend ratio. The sample means of the risk free rate, equity premium, and the price-dividend ratio are 0.8%, 5.8%, and 3.38, respectively, while the median values of these moments obtained from 10000 simulations are 1.2%, 3.3%, and 2.95, respectively. The model generates somewhat lower volatility for the risk free rate, equity premium and price-dividend ratio than what is observed in the data.

Finally, the model generates a size premium of 8.3%, almost identical to the 9.4% value in the data, and a value premium of 3.7%, that is within the 95% confidence interval of the value 7.3% in the data.[6]

[5] Standard errors are Newey-West (1987) corrected using 2 lags.

[6] We do not take a stand on the specification of the dividend growth processes for the "Small",

5 Economic Interpretation of the Two Regimes

The regimes are correlated with the business cycle. In Figure 3, we plot the time-series of the probability, p_t, that the economy is in the first regime over $1930 - 2008$. The shaded areas mark recession years, defined here as years with two or more quarters in NBER-designated recession. The correlation between the probability series and a dummy variable that takes the value one in a recession year and zero otherwise is -0.42. Conditional on lower than 50% probability that the economy is in the first regime ($p_t < 0.5$), the probability of a recession in that year is 44.4%; conditional on higher than 50% probability that the economy is in the first regime ($p_t > 0.5$), the probability of a recession in that year is 8.2%. The association of the second regime with recessions is consistent with our earlier finding that the second regime is associated with lower persistence and higher volatility of the predictable component of consumption growth and has a shorter duration compared to the first regime.

— Figure 3 about here —

The regimes are also correlated with major stock market downturns. In Figure 3, the vertical dashed lines mark major stock market downturns, as defined in Barro and Ursua (2009). The correlation between the probability series and a dummy variable that takes the value one in years with a stock market downturn and zero otherwise is -0.40. Conditional on lower than 50% probability that the economy is in the first regime, the probability of a stock market downturn in that year is 44.4%; conditional on higher than 50% probability that the economy is in the first regime, the probability of a stock market downturn in that year is 8.2%. Note that the second regime does double duty by capturing both economic recessions and periods of stock market downturns. This rendition is necessarily imperfect because economic recessions and stock market downturns are related but distinct economic phenomena.[7]

In Table II, we report the annual sample mean and volatility, along with the associated asymptotic standard errors in parentheses, of the consumption, dividend, and GDP growth rates, the rate of inflation, the market-wide price-dividend ratio, risk free rate, market return, and equity, size, and value premia. In Panel A, we present these summary statistics for the 61 years over the period $1930 - 2008$ in which the probability that the economy is in the first regime exceeds 50%. In Panel B, we present these

"Large", "Growth", and "Value" portfolios. Therefore, the returns on these portfolios cannot be simulated. The model-implied value for the size premium is computed as follows:

$$E\left(R_s - R_b\right) = -\frac{\widehat{Cov}\left(R_{s,t} - R_{b,t}, \widehat{M}_t\right)}{\widehat{E}\left(\widehat{M}_t\right)},$$

where $\widehat{M}_t$ denotes the estimated time series of the pricing kernel, and the covariance and expectation operators are estimated using their sample analogs. A similar procedure is used to compute the model-implied value premium.

[7] The correlation between a dummy variable that takes the value one in a recession year and zero otherwise and a dummy variable that takes the value one in years with a stock market downturn and zero otherwise over the period $1930 - 2008$ is 0.45.

summary statistics for the 18 years in which the probability that the economy is in the first regime is below 50%. Given the small size of these subsamples, particularly the second one, the standard errors are large and differences in the point estimates across the two regimes are often statistically insignificant. However, the equity, size and value premia are much higher in the second regime than the first one. With the exception of the price-dividend ratio, the volatility of all variables is higher in the second regime than the first one.

— Table II about here —

In Table II, we also report the model-simulated median and 95% confidence interval (in square brackets) for the mean and volatility of the consumption and dividend growth rates, the log price-dividend ratio, risk free rate, market return, and equity premium. Consistent with the sample moments, the simulated equity premium is higher in the second regime than the first one; and the simulated volatility of all variables is higher in the second regime than the first one.

In Table III, we report regression coefficients (standard errors in parentheses) and (adjusted) $\overline{R}^2$ of in-sample linear regressions of the consumption and dividend growth rates and the market return and equity premium on the log price-dividend ratio as predictive variable in the two regimes. In the first regime, the market return and equity premium are more predictable than in the second one: the price-dividend ratio has coefficients with the right sign in the first regime and $\overline{R}^2$ of 1.4% and 1.5%, respectively, while it performs poorly at forecasting the market return and premium in the second regime with $\overline{R}^2$ of -5.9% and -1.3%, respectively. In the second regime, the consumption and dividend growth rates are much more predictable than in the first one. The price-dividend ratio has a statistically significant coefficient in the second regime for the consumption and dividend growth rates and $\overline{R}^2$ of 41.3% and 23.7%, respectively, but it performs poorly at forecasting the growth rates in the first regime with $\overline{R}^2$ of -1.0% and -1.6%, respectively. These results should be interpreted with caution because the second regime has only 18 observations. The differences in predictability across regimes may shed light on why the empirical evidence on predictability which does not explicitly account for regime shifts is not robust in subperiods and its interpretation is controversial; and why recognition of structural breaks has important implications for return predictability (Lettau and Van Nieuwerburgh (2008), Pastor and Stambaugh (2001), and Paye and Timmermann (2006)).

— Table III about here —

In Figure 4, we plot the time-series of the second state variable, x_t, over $1930-2008$. The model implies that the expected consumption and dividend growth rates are affine functions of the two state variables, x_t and p_t, (equations (22) and (23)). We show in Section 7 that x_t has strong in-sample and out-of-sample forecasting power for the aggegate consumption and dividend growth rates.

— Figure 4 about here —

6 Forecasting the Equity, Size, and Value Premia

We perform forecasting regressions of the equity, size, and value premia on the model state variables and compare the results with those obtained from corresponding regressions of the premia on the market-wide price-dividend ratio and risk free rate. In Section 6.1, we estimate the model parameters over the period $1930-2009$, extract the time series of the state variables, and perform in-sample forecasting regressions over the same period. In Section 6.2, we estimate the model parameters over the subperiod $1930-1975$, extract the time series of the state variables, and perform in-sample forecasting regressions and out-of-sample predictive regressions over the non-overlapping subperiod $1976-2009$. The results provide strong evidence in favor of the model, with the out-of-sample predictive regressions providing the strongest evidence.

6.1 In-Sample Forecasting: 1930-2009

The expected equity premium implied by the model is an affine function of the two state variables and their product (equation (21)). We estimate the model parameters over the period $1930-2009$ and extract the time series of the state variables, as described in Section 4. We then perform an in-sample forecasting regression of the realized equity premium on the lagged state variables and their product. The results are displayed in the first row of Table IV, Panel A. The coefficient on p is statistically significant and the $\overline{R}^2$ is 7.4%. The performance of the model is superior to the forecasting performance of the price-dividend ratio with $\overline{R}^2$ 2.8% (second row); the joint forecasting performance of the price-dividend ratio and risk free rate with $\overline{R}^2$ 3.8% (third row); and the joint forecasting performance of the price-dividend ratio, risk free rate, and their product, with $\overline{R}^2$ 3.4% (fourth row). The regression in the third row is performed to facilitate comparison with the single-regime Bansal and Yaron (2004) model that implies that the expected equity premium is an affine function of the lagged price-dividend ratio and risk free rate.[8] The regression in the fourth row is motivated by the model implication that the two state variables are non-linear functions of the aggregate log price-dividend ratio and risk free rate ((see equations (18) and (19))).

— Table IV about here —

The sample size premium (9.4%) and value premium (7.3%) provide an alternative channel to examine the empirical plausibility of the model. We perform an in-sample forecasting regression of the realized size premium on the lagged state variables and their product. The results are displayed in the first row of Table IV, Panel B. The coefficients on p and the product, xp, are strongly statistically significant and the $\overline{R}^2$ is 14.3%. The performance of the model is superior to the forecasting performance of the price-dividend ratio with $\overline{R}^2$ 0.6% (second row); the joint forecasting performance of the price-dividend ratio and risk free rate with $\overline{R}^2$ 2.8% (third row); and the joint forecasting performance of the price-dividend ratio, risk free rate, and their product,

[8] See Constantinides and Ghosh (2011) for a derivation of this result.

with $\overline{R}^2$ 8.9% (fourth row). Finally we perform an in-sample forecasting regression of the realized value premium on the lagged state variables and their product. The results are displayed in the first row of Table IV, Panel C. The coefficients on x and p are strongly statistically significant and the $\overline{R}^2$ is 4.8%. The performance of the model is superior to the forecasting performance of the price-dividend ratio with $\overline{R}^2$ -0.7% (second row); the joint forecasting performance of the price-dividend ratio and risk free rate with $\overline{R}^2$ -2.0% (third row); and the joint forecasting performance of the price-dividend ratio, risk free rate, and their product, with $\overline{R}^2$ -1.6% (fourth row).

The results are illustrated in Figure 5. Panels A, C, and E display the realized equity, size, and value premia (black solid line), respectively, along with their predicted values from the forecasting regressions implied by the model (green dotted line), and linear forecasting regressions using the log price-dividend ratio as a predictor variable (red dashed line). The time series of the premia predicted by the model line up more closely with the actual realized time series compared to the time series predicted by the price-dividend ratio. Panels B, D, and F display the cumulative squared demeaned equity, size, and value premia, respectively, minus the cumulative squared regression residual from the alternative forecasting regression specifications: the forecasting regression implied by the model (black solid line) and a linear forecasting regression with the log price-dividend ratio as a predictor variable (red dashed line). The figure reveals the superior forecasting performance of the regime shifts model relative to the price-dividend ratio for the equity, size, and value premia.

— Figure 5 about here —

6.2 In-Sample Forecasting and out-of-Sample Prediction, 1976-2009

We re-examine the ability of the regime shifts model to forecast in sample over the subperiod $1976-2009$ for two reasons. First, it facilitates comparison with the extant literature that documents poor in-sample performance of forecasting models over this particular subperiod (see Welch and Goyal (2008)). Second, it allows us to estimate the model parameters over the first subperiod $1930-1975$ and examine the forecasting performance of the model over the non-overlapping second subperiod $1976-2009$, thereby eliminating the potential look-ahead bias introduced by estimating the model parameters over the same period over which we forecast the premia.

We also examine the ability of the regime shifts model to predict out of sample over the subperiod $1976-2009$ and compare our results to the extant literature over the same subperiod (see Welch and Goyal (2008)). At each year t, starting from 1975, we forecast the premia in the year $t+1$ as follows. First, we estimate the model parameters over the period $1930-1975$ and extract the time series of the state variables. This approach is conservative because we do not use all the information in the history from 1930 to time t in estimating the model parameters. Second, we estimate the coefficients of the lagged values of x, p, and xp from a regression over the period 1930 to time t and use these coefficients to forecast the premia at time $t+1$.[9] The out-of-sample

[9] Campbell and Thompson (2008) point out that the rolling out-of-sample predictive regressions are

performance of these forecasts is evaluated using an out-of-sample R^2 statistic as in Campbell and Thompson (2008) and Welch and Goyal (2008):

$$R^2_{OOS} = 1 - \frac{MSE_A}{MSE_N}, \tag{26}$$

where MSE_A denotes the mean-squared prediction error from the predictive regression implied by the model and MSE_N denotes the mean-squared prediction error of the historical average return. If R^2_{OOS} is positive, then the predictive regression has lower mean-squared prediction error than the historical average return.

The in-sample and out-of-sample results on the equity premium are reported in Table V, Panel A. The first row displays results of a forecasting regression with the state variables and their product as predictive variables. The $\overline{R}^2$ of the regression is 8.5%. The performance of the model is superior to the forecasting performance of the price-dividend ratio with $\overline{R}^2$ 2.2% (second row); superior to the joint forecasting performance of the price-dividend ratio and risk free rate with $\overline{R}^2$ -1.0% (third row); and comparable to the joint forecasting performance of the price-dividend ratio, risk free rate, and their product, with $\overline{R}^2$ 9.5% (fourth row). However, the regme shifts model retains its predictive performance out of sample with R^2_{OOS} 5.2%, while the price-dividend ratio, the combined price-dividend ratio and risk free rate, and the combined price-dividend ratio, risk free rate, and their product all have negative R^2_{OOS}.

— Table V about here —

Panel B displays results for the size premium. The first row shows that the in-sample forecasting regression with x, p, and their product as predictive variables yields an $\overline{R}^2$ of 25.5%. The model retains its predictive performance out of sample with R^2_{OOS} 22.6%. By contrast, the price-dividend ratio and the combined price-dividend ratio and risk free rate have negative R^2_{OOS} while the combined price-dividend ratio, risk free rate, and their product yields an R^2_{OOS} of 5.3%.

Panel C, displays results for the value premium. The first row shows that the in-sample forecasting regression with x, p, and their product as predictive variables yields an $\overline{R}^2$ of 5.9%. The model gives an R^2_{OOS} of 0%. The price-dividend ratio, the combined price-dividend ratio and risk free rate, and the combined price-dividend ratio, risk free rate, and their product yield negligible or negative R^2 both in sample and out of sample.

The in-sample and out-of-sample results are illustrated in Figures 6 and 7, respectively. The description of the figures is similar to that of Figure 5. The overall conclusion is that, over the subperiod $1976-2009$, the model forecasts in sample and

estimated over short sample periods, particularly at the beginning of the forecast evaluation period, and can, therefore, easily generate perverse results, such as a negative coefficient when theory suggests that the coefficient should be positive. In all of our out-of-sample predictive regressions for the equity premium, we impose two restrictions suggested in Campbell and Thompson (2008): a) we set the regression coefficients to zero whenever they have the "wrong" sign (different from the theoretically expected sign obtained from the model), and b) we assume that investors rule out a negative equity premium, and set the forecast to zero whenever it is negative.

predicts out of sample the equity, size, and value premia far better than the price-dividend ratio, the combined price-dividend ratio and risk free rate, and the combined price-dividend ratio, risk free rate, and their product.

— Figures 6 and 7 about here —

7 In-Sample Forecasting and out-of-Sample Prediction of Consumption and Dividend Growth

The model implies that the expected consumption growth rate is linear in x_t with coefficient one and linear in p_t with coefficient smaller than 0.01 (equation (22)). It also implies that the expected dividend growth rate is linear in x_t with coefficient $\phi = 3.5$ and linear in p_t with coefficient smaller than 0.01 (equation (23)). We show that the state variables forecast the consumption and dividend growth rates with the right sign and order of magnitude of the regression coefficients. The results are consistent with the presence of a predictable component of the consumption and dividend processes and the mechanism by which the state variables drive the dynamics. They are also consistent with the findings of the earlier literature on the price-dividend ratio as an unreliable predictor of consumption and dividend growth.

We estimate the model parameters over the period $1930 - 2009$, extract the time series of the state variables, and perform an in-sample linear forecasting regression of consumption growth on the two state variables over the same period. The results are reported in the first row of Table VI, Panel A. The $\overline{R}^2$ is 8.0% but the regression coefficient on x_t has the wrong sign. This is largely driven by the inability of the state variable x_t to explain the sharp movements in consumption during the prewar period, as shown below. The regression on the price-dividend ratio yields $\overline{R}^2$ 6.8% (second row) and on the price-dividend ratio and risk free rate yields $\overline{R}^2$ 8.4% (third row).

— Table VI about here —

We repeat the above forecasting regressions over the subperiod $1947-2009$, thereby avoiding the prewar period. The results are reported in Panel B. The forecasting regression of consumption growth rate on the two state variables yields statistically significant coefficient on x_t of the right sign and $\overline{R}^2$ 21.4%. The regression on the price-dividend ratio yields $\overline{R}^2$ 6.7% (second row) and on the price-dividend ratio and risk free rate yields $\overline{R}^2$ 23.9% (third row).

In Panel C, we report the results of in-sample forecasting regressions and out-of sample predictive regressions over the subperiod $1976-2009$. The forecasting regression of consumption growth rate on the two state variables yields statistically significant coefficient on x_t of the right sign. The $\overline{R}^2$ is 15.6% in sample and remains positive albeit small (0.7%) out of sample. The regression on the price-dividend ratio yields zero $\overline{R}^2$ in sample and large negative R^2_{OOS} out of sample (second row); and the regression on the price-dividend ratio and risk free rate yields $\overline{R}^2$ 23.6% in sample and large negative R^2_{OOS} out of sample (third row).

In Table VII, we report corresponding results for in-sample forecasting and out-of-sample prediction of the aggregate dividend growth rate. Over the period $1930-2009$, an in-sample linear forecasting regression of dividend growth on the two state variables yields $\overline{R}^2$ 11.7% but the coefficient on x_t is not statistically significant (first row, Panel A). The regression on the price-dividend ratio yields $\overline{R}^2$ 8.0% (second row) and a regression on the price-dividend ratio and risk free rate yields $\overline{R}^2$ 7.0% (third row).

— Table VII about here —

Over the subperiod $1947-2009$, an in-sample linear forecasting regression of dividend growth on the two state variables yields $\overline{R}^2$ 8.7% and positive and statistically significant regression coefficient on x_t (first row, Panel B). The regression on the price-dividend ratio yields negative $\overline{R}^2$ (second row) and on the price-dividend ratio and risk free rate yields $\overline{R}^2$ 8.4% (third row).

In Panel C, we report the results of in-sample forecasting regressions and out-of sample predictive regressions over the subperiod $1976-2009$. The forecasting regression of dividend growth rate on the two state variables yields a coefficient on x_t of the right sign. The $\overline{R}^2$ is 2.9% in sample and 4.4% out of sample. The regression on the price-dividend ratio yields zero $\overline{R}^2$ in sample and large negative R^2_{OOS} out of sample (second row); and the regression on the price-dividend ratio and risk free rate yields $\overline{R}^2$ 7.6% in sample and large negative R^2_{OOS} out of sample (third row).

8 Forecasting the Variance of the Market Return and Consumption and Dividend Growth

We estimate the conditional variance of the annual market return as the sum of squares of the twelve monthly log returns. In the first row of Table VIII, Panel A, we report the results of the in-sample forecasting regression of this conditional variance on the state variables and their product over $1930-2009$. The regression coefficient on p is statistically significant and the $\overline{R}^2$ is 2.1%. We also report results of in-sample forecasting regressions on the price-dividend ratio (Row 2), the price-dividend ratio and risk free rate (Row 3), and the price-dividend ratio, risk free rate, and their product (Row 4). None of the regression coefficients is statistically significant and the $\overline{R}^2$ varies from -0.9% to 1.0%.

— Table VIII about here —

The superior performance of the model in forecasting the conditional variance of the market return is illustrated in Figure 8 that plots the realized variance (black solid line) along with its predicted value from the forecasting regression implied by the regime shift model (green dotted line) and a linear forecasting regression using the market-wide price-dividend ratio as a predictor variable (red dashed line). Note that the time series of the variance predicted by the model lines up much more closely with the actual realized time series compared to the time series predicted by the price-dividend ratio.

— Figure 8 about here —

The model also implies that the conditional variance of consumption growth is linear in the state variable p_t (equation (24)). The conditional variance is computed as the squared residual from a regression of consumption growth on the two state variables. In the first row of Table VIII, Panel B, we report the results of the in-sample forecasting regression of this conditional variance over $1930-2009$ on the state variable p_t. The regression coefficient is strongly statistically significant with the right sign and the $\overline{R}^2$ is 12.7%. The in-sample forecasting regression on the price-dividend ratio yields a statistically insignificant coefficient and $\overline{R}^2$ 2.7% (Row 2). The in-sample forecasting regression on the price-dividend ratio and risk free rate yields a marginally significant coefficient for the price-dividend ratio and $\overline{R}^2$ of only 0.6% (Row 3).

Finally, the model implies that the conditional variance of dividend growth is linear in the state variable p_t (equation (25)). In the first row of Table VIII, Panel C, we report the results of the in-sample forecasting regression of this conditional variance over $1930-2009$ on the state variable p_t. The regression coefficient is statistically significant with the right sign and the $\overline{R}^2$ is 15.5%. The in-sample forecasting regression on the price-dividend ratio yields a statistically significant coefficient and $\overline{R}^2$ 3.4% (Row 2). The in-sample forecasting regression on the price-dividend ratio and risk free rate yields a significant coefficient only for the price-dividend ratio and $\overline{R}^2$ 2.4% (Row 3).

9 Concluding Remarks

We present an exchange economy with consumption and dividend processes that differ across two regimes and derive the equilibrium implications on the stochastic discount factor, the price of the dividend claim, and the risk free rate. At the estimated parameter values, the model implies that the second regime is shorter in duration than the first one, the expected consumption and dividend growth rates are less persistent and more volatile in the second regime compared to the first one, and that consumption and dividend growth, the return on the market, and the risk free rate are more volatile in the second regime than the first one. We verify these predictions over the period $1930-2009$. The second regime is associated with recessions and market downturns; and consumption and dividend growth, the return on the market, and the risk free rate are more volatile in the second regime than the first one.

The model further implies that the conditional mean of the consumption and dividend growth, the market return, and the equity premium differ across regimes. We show that the model-implied state variables perform significantly better at in-sample forecasting and out-of-sample prediction of the equity, size, and value premia, and the aggregate consumption and dividend growth rates than linear regressions with the price-dividend ratio and risk free rate as predictive variables.

High on our agenda is the application of the model to explain of the cross-section of equity, bond, and derivative returns. Also high on our agenda is the investigation on the number of regimes that are needed to adequately describe the economy. At present, our second regime does double duty by capturing both economic recessions and market downturns. This rendition is necessarily imperfect because economic recessions and market downturns are related but distinct economic phenomena. The challenge is the judicious increase of the number of regimes in a model that retains computational and empirical tractability.

References

[1] Ang, A. and Bekaert, G., (2007), “Stock Return Predictability: Is it There?” *The Review of Financial Studies*, 20(3), 651-707.

[2] Baker, M. and J. Wurgler, (2006), "Investor Sentiment and the Cross-Section of Stock Returns," *Journal of Finance*, 61, 1645-1680.

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

[4] Bansal, R., and Shaliastovich, I., (2011), "Learning and Asset Price Jumps," *Review of Financial Studies*, Forthcoming.

[5] van Binsbergen, J. H., and Koijen, R. S. J., (2010), "Predictive Regressions: A Present-Value Approach," *Journal of Finance*, 65, 1439-1471.

[6] Boudoukh, J., Richardson, M., and Whitelaw R. (2008) "The Myth of Long Horizon Predictability," *Review of Financial Studies*, 24, 1577-1605.

[7] Brandt, M. W., and Kang, Q., (2004), "On the Relation Between the Conditional Mean and Volatility of Stock Returns: A Latent VAR Approach," *Journal of Financial Economics*, 72, 217-257.

[8] 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.

[9] Campbell, J. Y., and Thompson, S., (2008), "Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?," *Review of Financial Studies*, 21(4), 1509-1531.

[10] Cochrane, J. H., (2008), "The Dog That Did Not Bark: A Defense of Return Predictability," *Review of Financial Studies*, 21(4). 1533-1575.

[11] Constantinides, G. M., and Ghosh, A., (2011), "Asset Pricing Tests with Long Run Risks in Consumption Growth," *Working paper.*

[12] Constantinides, G. M., Jackwerth, J. C., and Savov, A. Z., (2011), “The Puzzle of Index Option returns,” *Working paper*, University of Chicago.

[13] Dow, C. H., (1920), *Scientific Stock Speculation*. The Magazine of Wall Street.

[14] Drechsler, I., (2009), "Uncertainty, Time-Varying Fear, and Asset Prices," *Working paper.*

[15] 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.

[16] Fama, E. F. and French, K. R. (1988), "Dividend Yields and Expected Stock Returns," *Journal of Financial Economics*, 22, 3-25.

[17] 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.

[18] Hall, R. E., (1978), "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence," *Journal of Political Economy*, 86(6), 971-987.

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

[20] Hore, S., (2010), "Equilibrium Predictability and Other Return Characteristics, " *Working paper.*

[21] Kelly, B., and Pruitt, S., (2010), "Disaggregate Valuation Ratios and Market Expectations," *Working Paper*, University of Chicago.

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

[23] Lettau, M., and Ludvigson, S., (2001), "Consumption, Aggregate Wealth, and Expected Stock Returns," *Journal of Finance*, 56, 815-849.

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

[25] Menzly, L., Santos, T., and Veronesi, P., (2004), "Understanding Predictability," *Journal of Political Economy*, 112(1), 1-47.

[26] Newey, W. K., and West, K. D., (1987), "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 702-708.

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

[28] Pastor, L., and Stambaugh, R. F., (2009), "Predictive Systems: Living with Imperfect Predictors," *Journal of Finance*, 64, 1583–1628.

[29] Paye, B., and Timmerman, A., (2006), "Instability of Return Prediction Models," *Journal of Empirical Finance*, 13 (3), 274-315.

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

[31] Welch, I., and Goyal, A., (2008), "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction," *Review of Financial Studies*, 21(4), 1455-1508.

A Appendix

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

A.1 Restriction on $\varepsilon(1)$

The law of motion of the probability, p_t, is

$$p_{t+1} = f(p_t) + \varepsilon_{t+1}, \tag{27}$$

where

$$\begin{aligned} f(p_t) &\equiv \pi_1 p_t + (1-\pi_2)(1-p_t) \\ &= (1-\pi_2) + (\pi_1+\pi_2-1)\, p_t. \end{aligned}$$

Case 1: $\pi_1 + \pi_2 - 1 > 0$

In this case, $f(p_t)$ is a monotonically increasing function of p_t. Since $p_t \in [0,1]$, we have

$$f(p_t) \in [1-\pi_2, \pi_1]\,. \tag{28}$$

Given this range for $f(p_t)$, equation (27) implies the following restriction on ε_{t+1} so as to keep p_{t+1} in its permissible range, i.e. in the unit interval:

$$\varepsilon_{t+1} \in [-(1-\pi_2), 1-\pi_1]\,. \tag{29}$$

In Section 2.1, we recognize that the mean of the residual ε_{t+1} conditional on the regime at time $t+1$, may differ from its unconditional value of zero:

$$E(\varepsilon_{t+1}|s_{t+1}=i) = \varepsilon(i)\,, \quad i = 1,2.$$

Since the unconditional expectation of ε_{t+1} is zero, the law of iterated expectations implies

$$\frac{1-\pi_2}{2-\pi_1-\pi_2}\varepsilon(1) + \left[1 - \frac{1-\pi_2}{2-\pi_1-\pi_2}\right]\varepsilon(2) = 0,$$

which implies

$$\varepsilon(2) = -\varepsilon(1)\,\frac{1-\pi_2}{1-\pi_1}. \tag{30}$$

Now, from equation (28), $f(p_t)+\varepsilon(1) \le 1$ provided $\varepsilon(1) \le (1-\pi_1)$. This condition is satisfied by equation (29).

We also require $f(p_t) + \varepsilon(2) \le 1$, i.e.

$$f(p_t) - \varepsilon(1)\,\frac{1-\pi_2}{1-\pi_1} \le 1.$$

Since the above restriction must hold for all values of p_t and since the maximum possible value of $f(p_t)$ is π_1 (equation (28)), we have

$$\pi_1 - \varepsilon(1)\frac{1-\pi_2}{1-\pi_1} \leq 1,$$

which implies

$$\varepsilon(1) \geq -\frac{(1-\pi_1)^2}{1-\pi_2}. \tag{31}$$

Now, from equation (28), $f(p_t) + \varepsilon(1) \geq 0$ provided $\varepsilon(1) \geq -(1-\pi_2)$. This condition is satisfied by equation (29).

Finally, we require $f(p_t) + \varepsilon(2) \geq 0$, i.e.

$$f(p_t) - \varepsilon(1)\frac{1-\pi_2}{1-\pi_1} \geq 0.$$

Since the above restriction must hold for all values of p_t and since the minimum possible value of $f(p_t)$ is $(1-\pi_1)$ (equation (28)), we have

$$(1-\pi_1) - \varepsilon(1)\frac{1-\pi_2}{1-\pi_1} \geq 0,$$

which implies $\varepsilon(1) \leq 1-\pi_1$. This condition is satisfied by equation (29).

Therefore, from equations (29) and (31), the pemissible range for $\varepsilon(1)$ is

$$\varepsilon(1) \in \left[\max\left(-(1-\pi_2), -\frac{(1-\pi_1)^2}{1-\pi_2}\right), 1-\pi_1\right].$$

Case 2: $\pi_1 + \pi_2 - 1 < 0$

In this case, $f(p_t)$ is a monotonically decreasing function of p_t. Since $p_t \in [0,1]$, we have

$$f(p_t) \in [\pi_1, 1-\pi_2], \tag{32}$$

$$\varepsilon_{t+1} \in [-\pi_1, \pi_2]. \tag{33}$$

Now, from equation (32), $f(p_t) + \varepsilon(1) \leq 1$ provided $\varepsilon(1) \leq \pi_2$. This condition is satisfied by equation (33).

We also require $f(p_t) + \varepsilon(2) \leq 1$, i.e.

$$f(p_t) - \varepsilon(1)\frac{1-\pi_2}{1-\pi_1} \leq 1.$$

Since the above restriction must hold for all values of p_t and since the maximum possible value of $f(p_t)$ is $1-\pi_2$ (equation (32)), we have

$$1-\pi_2 - \varepsilon(1)\frac{1-\pi_2}{1-\pi_1} \leq 1,$$

which implies

$$\varepsilon(1) \geq -\frac{(1-\pi_1)\,\pi_2}{1-\pi_2}. \tag{34}$$

Now, from equation (32), $f(p_t) + \varepsilon(1) \geq 0$ provided $\varepsilon(1) \geq -\pi_1$. This condition is satisfied by equation (33).

Finally, we require $f(p_t) + \varepsilon(2) \geq 0$, i.e.

$$f(p_t) - \varepsilon(1)\frac{1-\pi_2}{1-\pi_1} \geq 0.$$

Since the above restriction must hold for all values of p_t and since the minimum possible value of $f(p_t)$ is π_1 (equation (32)), we have

$$\pi_1 - \varepsilon(1)\frac{1-\pi_2}{1-\pi_1} \geq 0,$$

which implies

$$\varepsilon(1) \leq \frac{(1-\pi_1)\,\pi_1}{1-\pi_2}. \tag{35}$$

Therefore, from equations (33), (34) and (35), the pemissible range for $\varepsilon(1)$ is

$$\varepsilon(1) \in \left[\max\left(-\pi_1, -\frac{(1-\pi_1)\,\pi_2}{1-\pi_2}\right), \min\left(\pi_2, \frac{(1-\pi_1)\,\pi_1}{1-\pi_2}\right)\right].$$

A.2 Derivation of Pricing Restrictions

Note that Assumptions (8) to (9) imply the following results:

i)

$$\begin{aligned} & E\left[\sigma_{s_{t+1}} e_{t+1} | F(t)\right] \\ = \; & f(p_t)\sigma_1 E\left[e_{t+1}|F(t), s_{t+1}=1\right] + (1-f(p_t))\,\sigma_2 E\left[e_{t+1}|F(t), s_{t+1}=2\right], \\ = \; & (\sigma_1-\sigma_2)\,e(1) f(p_t), \end{aligned} \tag{36}$$

where the first equality follows from the law of iterated expectations, and the second equality follows since

$$\begin{aligned} & f(p_t)E\left[e_{t+1}|F(t), s_{t+1}=1\right] + (1-f(p_t))\,E\left[e_{t+1}|F(t), s_{t+1}=2\right] \\ = \; & E\left[e_{t+1}|F(t)\right], \\ = \; & 0. \end{aligned}$$

ii)

$$
\begin{aligned}
& E\left[\sigma_{s_{t+1}}\eta_{t+1}|F(t)\right] \\
= \; & f(p_t)\sigma_1 E\left[\eta_{t+1}|F(t), s_{t+1}=1\right] + (1-f(p_t))\,\sigma_2 E\left[\eta_{t+1}|F(t), s_{t+1}=2\right], \\
= \; & (\sigma_1-\sigma_2)\,\eta(1) f(p_t), \qquad (37)
\end{aligned}
$$

where the first equality follows from the law of iterated expectations, and the second equality follows since

$$
\begin{aligned}
& f(p_t)E\left[\eta_{t+1}|F(t), s_{t+1}=1\right] + (1-f(p_t))\,E\left[\eta_{t+1}|F(t), s_{t+1}=2\right] \\
= \; & E\left[\eta_{t+1}|F(t)\right], \\
= \; & 0.
\end{aligned}
$$

iii)

$$
\begin{aligned}
& E\left[\rho_{s_{t+1}}\varepsilon_{t+1}|F(t)\right] \\
= \; & f(p_t)\rho_1 E\left[\varepsilon_{t+1}|F(t), s_{t+1}=1\right] + (1-f(p_t))\,\rho_2 E\left[\varepsilon_{t+1}|F(t), s_{t+1}=2\right], \\
= \; & (\rho_1-\rho_2)\,\varepsilon(1) f(p_t), \qquad (38)
\end{aligned}
$$

where the first equality follows from the law of iterated expectations, and the second equality follows since

$$
\begin{aligned}
& f(p_t)E\left[\varepsilon_{t+1}|F(t), s_{t+1}=1\right] + (1-f(p_t))\,E\left[\varepsilon_{t+1}|F(t), s_{t+1}=2\right] \\
= \; & E\left[\varepsilon_{t+1}|F(t)\right], \\
= \; & 0.
\end{aligned}
$$

iv)

$$
\begin{aligned}
& E\left[\varepsilon_{t+1}\sigma_{s_{t+1}}e_{t+1}|F(t)\right] \\
= \; & f(p_t)\sigma_1 E\left[\varepsilon_{t+1}e_{t+1}|F(t), s_{t+1}=1\right] + (1-f(p_t))\,\sigma_2 E\left[\varepsilon_{t+1}e_{t+1}|F(t), s_{t+1}=2\right], \\
= \; & \sigma_{\varepsilon,e}\left\{\sigma_1 f(p_t) + \sigma_2\left(1-f(p_t)\right)\right\}, \qquad (39)
\end{aligned}
$$

where the first equality follows from the law of iterated expectations, and the second equality follows from equation (9).

A.2.1 Consumption Claim

We rely on the log-linear approximation for the continuous return on the consumption claim, $r_{c,t+1}$,

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

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,

$$z_t = p_t\left[A_0(1) + A_1(1)x_t\right] + (1-p_t)\left[A_0(2) + A_1(2)x_t\right].$$

The Euler equation for the consumption claim is,

$$E\left[\exp\left(m_{t+1} + r_{c,t+1}\right)|\digamma(t)\right] = 1, \tag{40}$$

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

Substituting the above expression for m_{t+1} into (40), we have,

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

which implies

$$E\left[\exp\left(\theta\log\delta - \frac{\theta}{\psi}\Delta c_{t+1} + \theta\left(\kappa_0 + \kappa_1 z_{t+1} - z_t + \Delta c_{t+1}\right)\right)|\digamma(t)\right] = 1$$

By Taylor series expansion up to quadratic terms, we obtain the following:

$$\begin{aligned}
& \theta\log\delta + \theta\left(\kappa_0 - z_t\right) + \left(\theta - \frac{\theta}{\psi}\right)E\left[\Delta c_{t+1}|\digamma(t)\right] + \theta\kappa_1 E\left[z_{t+1}|\digamma(t)\right] \\
& + \frac{1}{2}var\left(\left(\theta - \frac{\theta}{\psi}\right)\Delta c_{t+1} + \theta\kappa_1 z_{t+1}|\digamma(t)\right) \\
= \; & 0,
\end{aligned}$$

which implies,

$$\begin{aligned}
& \theta\log\delta + \theta\kappa_0 - \theta\left\{p_t\left[A_0(1) + A_1(1)x_t\right] + (1-p_t)\left[A_0(2) + A_1(2)x_t\right]\right\} \\
& + \left(\theta - \frac{\theta}{\psi}\right)E\left[\Delta c_{t+1}|\digamma(t)\right] + \theta\kappa_1 E\left[z_{t+1}|\digamma(t)\right] + \frac{1}{2}var\left(\left(\theta - \frac{\theta}{\psi}\right)\Delta c_{t+1} + \theta\kappa_1 z_{t+1}|\digamma(t)\right) \\
= \; & 0
\end{aligned}$$

We approximate the conditional variance, $var\left(\left(\theta - \frac{\theta}{\psi}\right)\Delta c_{t+1} + \theta\kappa_1 z_{t+1}|\digamma(t)\right)$, with the constant, Σ, and write the above equation as

$$\begin{aligned}
& \theta\log\delta + \theta\kappa_0 - \theta\left\{p_t\left[A_0(1) + A_1(1)x_t\right] + (1-p_t)\left[A_0(2) + A_1(2)x_t\right]\right\} \\
& + \left(\theta - \frac{\theta}{\psi}\right)E\left[\Delta c_{t+1}|\digamma(t)\right] + \theta\kappa_1 E\left[z_{t+1}|\digamma(t)\right] + \frac{1}{2}\Sigma \\
= \; & 0
\end{aligned}$$

The parameter Σ is a function of the deeper parameters of the joint distribution of the error terms e_{t+1}, η_{t+1}, u_{t+1}, ε_{t+1}, and s_{t+1}(e.g. $E\left[e_{t+1}\eta_{t+1}|\mathcal{F}(t), s_{t+1}\right]$, $E\left[\varepsilon_{t+1}\eta_{t+1}|\mathcal{F}(t), s_{t+1}\right]$, $E\left[\varepsilon_{t+1}e_{t+1}\eta_{t+1}|\mathcal{F}(t), s_{t+1}\right]$). In our empirical work, we treat Σ as a free parameter.

We calculate $E\left[\Delta c_{t+1}|\mathcal{F}(t)\right]$ as follows:

$$
\begin{aligned}
& E\left[\Delta c_{t+1}|\mathcal{F}(t)\right] \\
= \ & \mu + x_t + E\left[\sigma_{s_{t+1}}\eta_{t+1}|\mathcal{F}(t)\right] \\
= \ & \mu + x_t + \left(\sigma_1 - \sigma_2\right)\eta(1) f(p_t),
\end{aligned}
$$

where the second equality follows from equation (37).

We calculate $E\left[z_{t+1}|\mathcal{F}(t)\right]$ as follows:

$$
\begin{aligned}
& E\left[z_{t+1}|\mathcal{F}(t)\right] \\
= \ & E\left[\left(\begin{array}{c} \left(f(p_t) + \varepsilon_{t+1}\right)\left\{A_0(1) + A_1(1)\left(\rho_{s_{t+1}}x_t + \varphi_e\sigma_{s_{t+1}}e_{t+1}\right)\right\} + \\ \left(1 - f(p_t) - \varepsilon_{t+1}\right)\left\{A_0(2) + A_1(2)\left(\rho_{s_{t+1}}x_t + \varphi_e\sigma_{s_{t+1}}e_{t+1}\right)\right\} \end{array}\right)|\mathcal{F}(t)\right] \\
= \ & f(p_t)\left[A_0(1) + A_1(1)\left\{x_t\left(f(p_t)\rho_1 + \left(1 - f(p_t)\right)\rho_2\right) + \varphi_e E\left[\sigma_{s_{t+1}}e_{t+1}|\mathcal{F}(t)\right]\right\}\right] \\
& + A_0(1)E\left[\varepsilon_{t+1}|\mathcal{F}(t)\right] \\
& + A_1(1)x_t E\left[\rho_{s_{t+1}}\varepsilon_{t+1}|\mathcal{F}(t)\right] \\
& + A_1(1)\varphi_e E\left[\varepsilon_{t+1}\sigma_{s_{t+1}}e_{t+1}|\mathcal{F}(t)\right] \\
& + \left(1 - f(p_t)\right)\left[A_0(2) + A_1(2)\left\{x_t\left(f(p_t)\rho_1 + \left(1 - f(p_t)\right)\rho_2\right) + \varphi_e E\left[\sigma_{s_{t+1}}e_{t+1}|\mathcal{F}(t)\right]\right\}\right] \\
& - A_0(2)E\left[\varepsilon_{t+1}|\mathcal{F}(t)\right] \\
& - A_1(2)x_t E\left[\rho_{s_{t+1}}\varepsilon_{t+1}|\mathcal{F}(t)\right] \\
& - A_1(2)\varphi_e E\left[\varepsilon_{t+1}\sigma_{s_{t+1}}e_{t+1}|\mathcal{F}(t)\right]
\end{aligned}
$$

We use equations (36), (38), and (39) to simplify the above expression as follows:

$$
\begin{aligned}
& E\left[z_{t+1}|\mathcal{F}(t)\right] \\
= \ & f(p_t)\left[A_0(1) + A_1(1)\left\{x_t\left(f(p_t)\rho_1 + \left(1 - f(p_t)\right)\rho_2\right)\right\}\right] \\
& + \left(1 - f(p_t)\right)\left[A_0(2) + A_1(2)\left\{x_t\left(f(p_t)\rho_1 + \left(1 - f(p_t)\right)\rho_2\right)\right\}\right] \\
& + \varphi_e\left(\sigma_1 - \sigma_2\right)e(1)\left\{A_1(1)f(p_t) + A_1(2)\left(1 - f(p_t)\right)\right\}f(p_t) \\
& + \left(A_1(1) - A_1(2)\right)\left(\rho_1 - \rho_2\right)\varepsilon(1)x_t f(p_t) \\
& + \left(A_1(1) - A_1(2)\right)\varphi_e\sigma_{\varepsilon,e}\left\{\sigma_1 f(p_t) + \sigma_2\left(1 - f(p_t)\right)\right\} \\
= \ & \left[A_0(1) - A_0(2) + \left(A_1(1) - A_1(2)\right)\rho_2 x_t\right]f(p_t) \\
& + A_0(2) + A_1(2)x_t\left(f(p_t)\rho_1 + \left(1 - f(p_t)\right)\rho_2\right) \\
& + \left(A_1(1) - A_1(2)\right)\left(\rho_1 - \rho_2\right)\varepsilon(1)x_t f(p_t) \\
& + \left(A_1(1) - A_1(2)\right)\varphi_e\sigma_{\varepsilon,e}\left\{\sigma_1 f(p_t) + \sigma_2\left(1 - f(p_t)\right)\right\} \\
& + \left(A_1(1) - A_1(2)\right)\left[\left(\rho_1 - \rho_2\right)x_t + \varphi_e\left(\sigma_1 - \sigma_2\right)e(1)\right]\left\{f(p_t)\right\}^2 \\
& + \varphi_e\left(\sigma_1 - \sigma_2\right)e(1)A_1(2)f(p_t)
\end{aligned}
$$

Finally, we write the Euler equation as

$$\begin{aligned}
& \theta \log \delta + \theta \kappa_0 - \theta \left\{ p_t \left[A_0(1) + A_1(2) x_t \right] + (1 - p_t) \left[A_0(2) + A_1(2) x_t \right] \right\} \\
& + \left(\theta - \frac{\theta}{\psi} \right) \left(\mu + x_t + (\sigma_1 - \sigma_2) \eta(1) f(p_t) \right) + \frac{1}{2} \Sigma \\
& + \theta \kappa_1 \begin{pmatrix} \left[A_0(1) - A_0(2) + (A_1(1) - A_1(2)) \rho_2 x_t \right] f(p_t) \\ + A_0(2) + A_1(2) x_t \left(f(p_t) \rho_1 + (1 - f(p_t)) \rho_2 \right) \\ + (A_1(1) - A_1(2)) (\rho_1 - \rho_2) \varepsilon(1) x_t f(p_t) \\ + (A_1(1) - A_1(2)) \varphi_e \sigma_{\varepsilon,e} \left\{ \sigma_1 f(p_t) + \sigma_2 (1 - f(p_t)) \right\} \\ + (A_1(1) - A_1(2)) \left[(\rho_1 - \rho_2) x_t + \varphi_e (\sigma_1 - \sigma_2) e(1) \right] \left\{ f(p_t) \right\}^2 \\ + \varphi_e (\sigma_1 - \sigma_2) e(1) A_1(2) f(p_t) \end{pmatrix} \\
= \; & 0
\end{aligned}$$

Collecting terms, we obtain

$$\begin{aligned}
& \begin{pmatrix} \theta \log \delta + \theta \kappa_0 - \theta A_0(2) + \left(\theta - \frac{\theta}{\psi} \right) \left(\mu + (\sigma_1 - \sigma_2) \eta(1)(1 - \pi_2) \right) + \theta \kappa_1 A_0(2) \\ + \theta \kappa_1 A_1(2) \varphi_e (\sigma_1 - \sigma_2) e(1)(1 - \pi_2) + \theta \kappa_1 \left[A_0(1) - A_0(2) \right] (1 - \pi_2) \\ + \theta \kappa_1 \left[A_1(1) - A_1(2) \right] \varphi_e (\sigma_1 - \sigma_2) e(1)(1 - \pi_2)^2 \\ + \theta \kappa_1 \left[A_1(1) - A_1(2) \right] \varphi_e \sigma_{\varepsilon,e} \left((\sigma_1 - \sigma_2)(1 - \pi_2) + \sigma_2 \right) + \frac{1}{2} \Sigma \end{pmatrix} \\
& + \begin{pmatrix} -\theta A_1(2) + \left(\theta - \frac{\theta}{\psi} \right) + \theta \kappa_1 A_1(2) \left[(\rho_1 - \rho_2)(1 - \pi_2) + \rho_2 \right] \\ + \theta \kappa_1 \left[A_1(1) - A_1(2) \right] (\rho_1 - \rho_2)(1 - \pi_2) \varepsilon(1) \\ \theta \kappa_1 \left[A_1(1) - A_1(2) \right] \left[(\rho_1 - \rho_2)(1 - \pi_2)^2 + \rho_2 (1 - \pi_2) \right] \end{pmatrix} x_t \\
& + \begin{pmatrix} \theta \left[A_0(2) - A_0(1) \right] + \left(\theta - \frac{\theta}{\psi} \right) (\sigma_1 - \sigma_2) \eta(1)(\pi_1 + \pi_2 - 1) \\ + \theta \kappa_1 A_1(2) \varphi_e (\sigma_1 - \sigma_2) e(1)(\pi_1 + \pi_2 - 1) \\ + \theta \kappa_1 \left[A_0(1) - A_0(2) \right] (\pi_1 + \pi_2 - 1) \\ + 2 \theta \kappa_1 \left[A_1(1) - A_1(2) \right] \varphi_e (\sigma_1 - \sigma_2) e(1)(1 - \pi_2)(\pi_1 + \pi_2 - 1) \\ + \theta \kappa_1 \left[A_1(1) - A_1(2) \right] \varphi_e \sigma_{\varepsilon,e} (\sigma_1 - \sigma_2)(\pi_1 + \pi_2 - 1) \end{pmatrix} p_t \\
& + \begin{pmatrix} \theta \left[A_1(2) - A_1(1) \right] + \theta \kappa_1 A_1(2) (\rho_1 - \rho_2)(\pi_1 + \pi_2 - 1) \\ + \theta \kappa_1 \left[A_1(1) - A_1(2) \right] \left[2(1 - \pi_2)(\pi_1 + \pi_2 - 1)(\rho_1 - \rho_2) + \rho_2 (\pi_1 + \pi_2 - 1) \right] \\ + \theta \kappa_1 \left[A_1(1) - A_1(2) \right] (\rho_1 - \rho_2) \varepsilon(1)(\pi_1 + \pi_2 - 1) \end{pmatrix} p_t x_t \\
& + \theta \kappa_1 \left[A_1(1) - A_1(2) \right] (\pi_1 + \pi_2 - 1)^2 (\rho_1 - \rho_2) p_t^2 x_t \\
& + \theta \kappa_1 \left[A_1(1) - A_1(2) \right] \varphi_e (\sigma_1 - \sigma_2) e(1)(\pi_1 + \pi_2 - 1)^2 p_t^2 \\
= \; & 0
\end{aligned}$$

We approximate the above expression to order x_t, p_t, and $p_t x_t$. Therefore, we expand the term p_t^2 as a Taylor series to first order around the unconditional mean, $\overline{p}$, of p_t. We note that $\overline{p} = \frac{1-\pi_2}{2-\pi_1-\pi_2}$. We obtain the following:

$$\begin{aligned} p_t^2 &\approx \overline{p}^2 + 2\overline{p}\left(p_t - \overline{p}\right) \\ &= -\overline{p}^2 + 2\overline{p}p_t \end{aligned}$$

Since the Euler equation holds for all observable states (x_t, p_t), we obtain the following 4 parameter restrictions:

Constant:

$$\left(\begin{array}{c} \theta \log \delta + \theta\kappa_0 - \theta A_0(2) + \left(\theta - \frac{\theta}{\psi}\right)\left(\mu + \left(\sigma_1 - \sigma_2\right)\eta(1)(1-\pi_2)\right) + \theta\kappa_1 A_0(2) \\ +\theta\kappa_1 A_1(2)\varphi_e\left(\sigma_1 - \sigma_2\right)e(1)(1-\pi_2) + \theta\kappa_1\left[A_0(1) - A_0(2)\right](1-\pi_2) \\ +\theta\kappa_1\left[A_1(1) - A_1(2)\right]\varphi_e\left(\sigma_1 - \sigma_2\right)e(1)(1-\pi_2)^2 \\ +\theta\kappa_1\left[A_1(1) - A_1(2)\right]\varphi_e\sigma_{\varepsilon,e}\left(\left(\sigma_1 - \sigma_2\right)(1-\pi_2) + \sigma_2\right) + \frac{1}{2}\Sigma \\ -\theta\kappa_1\left[A_1(1) - A_1(2)\right]\varphi_e\left(\sigma_1 - \sigma_2\right)e(1)(\pi_1 + \pi_2 - 1)^2\overline{p}^2 \end{array} \right) = 0$$

Coefficient of x_t:

$$\left(\begin{array}{c} -\theta A_1(2) + \left(\theta - \frac{\theta}{\psi}\right) + \theta\kappa_1 A_1(2)\left[\left(\rho_1 - \rho_2\right)(1-\pi_2) + \rho_2\right] \\ +\theta\kappa_1\left[A_1(1) - A_1(2)\right]\left(\rho_1 - \rho_2\right)(1-\pi_2)\varepsilon(1) \\ \theta\kappa_1\left[A_1(1) - A_1(2)\right]\left[\left(\rho_1 - \rho_2\right)(1-\pi_2)^2 + \rho_2(1-\pi_2)\right] \\ -\theta\kappa_1\left[A_1(1) - A_1(2)\right](\pi_1 + \pi_2 - 1)^2\left(\rho_1 - \rho_2\right)\overline{p}^2 \end{array} \right) = 0$$

Coefficient of p_t:

$$\left(\begin{array}{c} \theta\left[A_0(2) - A_0(1)\right] + \left(\theta - \frac{\theta}{\psi}\right)\left(\sigma_1 - \sigma_2\right)\eta(1)(\pi_1 + \pi_2 - 1) \\ +\theta\kappa_1 A_1(2)\varphi_e\left(\sigma_1 - \sigma_2\right)e(1)(\pi_1 + \pi_2 - 1) \\ +\theta\kappa_1\left[A_0(1) - A_0(2)\right](\pi_1 + \pi_2 - 1) \\ +2\theta\kappa_1\left[A_1(1) - A_1(2)\right]\varphi_e\left(\sigma_1 - \sigma_2\right)e(1)(1-\pi_2)(\pi_1 + \pi_2 - 1) \\ +\theta\kappa_1\left[A_1(1) - A_1(2)\right]\varphi_e\sigma_{\varepsilon,e}\left(\sigma_1 - \sigma_2\right)(\pi_1 + \pi_2 - 1) \\ +2\theta\kappa_1\left[A_1(1) - A_1(2)\right]\varphi_e\left(\sigma_1 - \sigma_2\right)e(1)(\pi_1 + \pi_2 - 1)^2\overline{p} \end{array} \right) = 0$$

Coefficient of $p_t x_t$:

$$\left(\begin{array}{c} \theta\left[A_1(2) - A_1(1)\right] + \theta\kappa_1 A_1(2)\left(\rho_1 - \rho_2\right)(\pi_1 + \pi_2 - 1) \\ +\theta\kappa_1\left[A_1(1) - A_1(2)\right]\left[2(1-\pi_2)(\pi_1 + \pi_2 - 1)\left(\rho_1 - \rho_2\right) + \rho_2(\pi_1 + \pi_2 - 1)\right] \\ +\theta\kappa_1\left[A_1(1) - A_1(2)\right]\left(\rho_1 - \rho_2\right)\varepsilon(1)(\pi_1 + \pi_2 - 1) \\ +2\theta\kappa_1\left[A_1(1) - A_1(2)\right](\pi_1 + \pi_2 - 1)^2\left(\rho_1 - \rho_2\right)\overline{p} \end{array} \right) = 0$$

The 4 linear equations can be solved to obtain the 4 parameters $A_0(1)$, $A_0(2)$, $A_1(1)$, and $A_1(2)$.

A.2.2 Dividend Claim

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

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

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,

$$z_{m,t} = p_t \left[A_{0,m}(1) + A_{1,m}(1)x_t\right] + (1 - p_t)\left[A_{0,m}(2) + A_{1,m}(2)x_t\right].$$

The Euler equation for the dividend claim is,

$$E\left[\exp\left(m_{t+1} + r_{m,t+1}\right)|\digamma(t)\right] = 1. \tag{41}$$

Substituting the expression for m_{t+1} from (14) into (41), we have,

$$E\left[\exp\left(\theta\log\delta - \frac{\theta}{\psi}\Delta c_{t+1} + (\theta-1)r_{c,t+1} + r_{m,t+1}\right)|\digamma(t)\right] = 1, \tag{42}$$

which implies:

$$E\left[\exp\left(\begin{array}{c}\theta\log\delta - \frac{\theta}{\psi}\Delta c_{t+1} + (\theta-1)\left(\kappa_0 + \kappa_1 z_{t+1} - z_t + \Delta c_{t+1}\right) \\ +\kappa_{0,m} + \kappa_{1,m} z_{m,t+1} - z_{m,t} + \Delta d_{t+1}\end{array}\right)|\digamma(t)\right] = 1.$$

Simplifying the above expression gives:

$$E\left[\exp\left(\begin{array}{c}\theta\log\delta + \left(-\frac{\theta}{\psi} + \theta - 1\right)\Delta c_{t+1} + \Delta d_{t+1} + (\theta-1)\kappa_0 + \kappa_{0,m} \\ -(\theta-1)z_t - z_{m,t} + (\theta-1)\kappa_1 z_{t+1} + \kappa_{1,m} z_{m,t+1}\end{array}\right)|\digamma(t)\right] = 1.$$

Performing a Taylor series expansion upto quadratic terms gives,

$$\begin{aligned} & \theta\log\delta + (\theta-1)\kappa_0 + \kappa_{0,m} - (\theta-1)z_t - z_{m,t} + \left(-\frac{\theta}{\psi} + \theta - 1\right)E\left[\Delta c_{t+1}|\digamma(t)\right] \\ & +E\left[\Delta d_{t+1}|\digamma(t)\right] + (\theta-1)\kappa_1 E\left[z_{t+1}|\digamma(t)\right] + \kappa_{1,m} E\left[z_{m,t+1}|\digamma(t)\right] \\ & +\frac{1}{2}var\left(\left(-\frac{\theta}{\psi} + \theta - 1\right)\Delta c_{t+1} + \Delta d_{t+1} + (\theta-1)\kappa_1 z_{t+1} + \kappa_{1,m} z_{m,t+1}|\digamma(t)\right) \\ = \; & 0. \end{aligned}$$

We approximate the conditional variance, $var\left(\left(-\frac{\theta}{\psi} + \theta - 1\right)\Delta c_{t+1} + \Delta d_{t+1} + (\theta-1)\kappa_1 z_{t+1} + \kappa_{1,m} z_{m,t+1}|\digamma(t)\right)$, with the constant, Σ' and write the above equation as:

$$
\begin{aligned}
& \theta \log \delta + (\theta - 1)\kappa_0 + \kappa_{0,m} - (\theta - 1)z_t - z_{m,t} + \left(-\frac{\theta}{\psi} + \theta - 1 \right) E\left[\Delta c_{t+1} | \mathcal{F}(t)\right] \\
& + E\left[\Delta d_{t+1} | \mathcal{F}(t)\right] + (\theta - 1)\kappa_1 E\left[z_{t+1} | \mathcal{F}(t)\right] + \kappa_{1,m} E\left[z_{m,t+1} | \mathcal{F}(t)\right] + \frac{1}{2}\Sigma' \\
= \ & 0 \qquad (43)
\end{aligned}
$$

We calculate $E\left[\Delta d_{t+1} | \mathcal{F}(t)\right]$ as follows:

$$
\begin{aligned}
& E\left[\Delta d_{t+1} | \mathcal{F}(t)\right] \\
= \ & \mu_d + \phi x_t + \varphi_d E\left[\sigma_{s_{t+1}} u_{t+1} | \mathcal{F}(t)\right] \\
= \ & \mu_d + \phi x_t + \varphi_d \left(\sigma_1 - \sigma_2\right) u(1) f(p_t).
\end{aligned}
$$

Similar calculations as in Appendix A.2.1 give the following expression for $E\left[z_{m,t+1} | \mathcal{F}(t)\right]$:

$$
\begin{aligned}
& E\left[z_{m,t+1} | \mathcal{F}(t)\right] \\
= \ & \left[A_{0,m}(1) - A_{0,m}(2) + \left(A_{1,m}(1) - A_{1,m}(2)\right) \rho_2 x_t\right] f(p_t) \\
& + A_{0,m}(2) + A_{1,m}(2) x_t \left(f(p_t)\rho_1 + \left(1 - f(p_t)\right) \rho_2\right) \\
& + \left(A_{1,m}(1) - A_{1,m}(2)\right) \left(\rho_1 - \rho_2\right) \varepsilon(1) x_t f(p_t) \\
& + \left(A_{1,m}(1) - A_{1,m}(2)\right) \varphi_e \sigma_{\varepsilon,e} \left\{\sigma_1 f(p_t) + \sigma_2 \left(1 - f(p_t)\right)\right\} \\
& + \left(A_{1,m}(1) - A_{1,m}(2)\right) \left[\left(\rho_1 - \rho_2\right) x_t + \varphi_e \left(\sigma_1 - \sigma_2\right) e(1)\right] \left\{f(p_t)\right\}^2 \\
& + \varphi_e \left(\sigma_1 - \sigma_2\right) e(1) A_{1,m}(2) f(p_t).
\end{aligned}
$$

Therefore, the Euler equation (43) may be written as:

$$
\begin{aligned}
&\theta\log\delta+(\theta-1)\kappa_0+\kappa_{0,m}-(\theta-1)\left\{p_t\left[A_0(0)+A_1(0)x_t\right]+(1-p_t)\left[A_0(1)+A_1(1)x_t\right]\right\}\\
&-\left\{p_t\left[A_{0,m}(0)+A_{1,m}(0)x_t\right]+(1-p_t)\left[A_{0,m}(1)+A_{1,m}(1)x_t\right]\right\}\\
&+\left(-\frac{\theta}{\psi}+\theta-1\right)\left(\mu+x_t+(\sigma_1-\sigma_2)\,\eta(1)f(p_t)\right)+\mu_d+\phi x_t+\varphi_d\,(\sigma_1-\sigma_2)\,u(1)f(p_t)+\frac{1}{2}\Sigma'\\
&+(\theta-1)\kappa_1\begin{pmatrix}\left[A_0(1)-A_0(2)+(A_1(1)-A_1(2))\,\rho_2x_t\right]f(p_t)\\ +A_0(2)+A_1(2)x_t\left(f(p_t)\rho_1+(1-f(p_t))\,\rho_2\right)\\ +\left(A_1(1)-A_1(2)\right)(\rho_1-\rho_2)\,\varepsilon(1)x_tf(p_t)\\ +\left(A_1(1)-A_1(2)\right)\varphi_e\sigma_{\varepsilon,e}\left\{\sigma_1f(p_t)+\sigma_2\left(1-f(p_t)\right)\right\}\\ +\left(A_1(1)-A_1(2)\right)\left[(\rho_1-\rho_2)\,x_t+\varphi_e\,(\sigma_1-\sigma_2)\,e(1)\right]\left\{f(p_t)\right\}^2\\ +\varphi_e\,(\sigma_1-\sigma_2)\,e(1)A_1(2)f(p_t)\end{pmatrix}\\
&+\kappa_{1,m}\begin{pmatrix}\left[A_{0,m}(1)-A_{0,m}(2)+(A_{1,m}(1)-A_{1,m}(2))\,\rho_2x_t\right]f(p_t)\\ +A_{0,m}(2)+A_{1,m}(2)x_t\left(f(p_t)\rho_1+(1-f(p_t))\,\rho_2\right)\\ +\left(A_{1,m}(1)-A_{1,m}(2)\right)(\rho_1-\rho_2)\,\varepsilon(1)x_tf(p_t)\\ +\left(A_{1,m}(1)-A_{1,m}(2)\right)\varphi_e\sigma_{\varepsilon,e}\left\{\sigma_1f(p_t)+\sigma_2\left(1-f(p_t)\right)\right\}\\ +\left(A_{1,m}(1)-A_{1,m}(2)\right)\left[(\rho_1-\rho_2)\,x_t+\varphi_e\,(\sigma_1-\sigma_2)\,e(1)\right]\left\{f(p_t)\right\}^2\\ +\varphi_e\,(\sigma_1-\sigma_2)\,e(1)A_{1,m}(2)f(p_t)\end{pmatrix}\\
=\;&0
\end{aligned}
$$

Simplifying, we obtain:

$$
\begin{aligned}
&\left(\begin{array}{c}
\theta \log \delta+(\theta-1) \kappa_0+\kappa_{0, m}-(\theta-1) A_0(2)-A_{0, m}(2)+\frac{1}{2} \Sigma^{\prime} \\
\left(-\frac{\theta}{\psi}+\theta-1\right)\left(\mu+\left(\sigma_1-\sigma_2\right) \eta(1)(1-\pi_2)\right)+\mu_d+\varphi_d\left(\sigma_1-\sigma_2\right) u(1)(1-\pi_2) \\
+(\theta-1) \kappa_1 A_1(2) \varphi_e\left(\sigma_1-\sigma_2\right) e(1)(1-\pi_2)+(\theta-1) \kappa_1\left[A_0(1)-A_0(2)\right](1-\pi_2) \\
+(\theta-1) \kappa_1\left[A_1(1)-A_1(2)\right] \varphi_e\left(\sigma_1-\sigma_2\right) e(1)(1-\pi_2)^2+(\theta-1) \kappa_1 A_0(2) \\
+(\theta-1) \kappa_1\left[A_1(1)-A_1(2)\right] \varphi_e \sigma_{\varepsilon, e}\left[\left(\sigma_1-\sigma_2\right)(1-\pi_2)+\sigma_2\right] \\
\kappa_{1, m}\left[A_{0, m}(1)-A_{0, m}(2)\right](1-\pi_2)+\kappa_{1, m} A_{0, m}(2) \\
+\kappa_{1, m}\left[A_{1, m}(1)-A_{1, m}(2)\right] \varphi_e \sigma_{\varepsilon, e}\left[\left(\sigma_1-\sigma_2\right)(1-\pi_2)+\sigma_2\right] \\
+\kappa_{1, m}\left[A_{1, m}(1)-A_{1, m}(2)\right] \varphi_e\left(\sigma_1-\sigma_2\right) e(1)(1-\pi_2)^2 \\
+\kappa_{1, m} A_{1, m}(1) \varphi_e\left(\sigma_1-\sigma_2\right) e(1)(1-\pi_2)
\end{array}\right) \\
&+\left(\begin{array}{c}
-(\theta-1) A_1(2)-A_{1, m}(2)+\left(-\frac{\theta}{\psi}+\theta-1\right)+\phi \\
(\theta-1) \kappa_1 A_1(2)\left[\left(\rho_1-\rho_2\right)(1-\pi_2)+\rho_2\right] \\
+(\theta-1) \kappa_1\left[A_1(1)-A_1(2)\right]\left(\rho_1-\rho_2\right)(1-\pi_2) \varepsilon(1) \\
(\theta-1) \kappa_1\left[A_1(1)-A_1(2)\right]\left[\left(\rho_1-\rho_2\right)(1-\pi_2)^2+\rho_2(1-\pi_2)\right] \\
+\kappa_{1, m} A_{1, m}(2)\left[\left(\rho_1-\rho_2\right)(1-\pi_2)+\rho_2\right] \\
+\kappa_{1, m}\left[A_{1, m}(1)-A_{1, m}(2)\right]\left(\rho_1-\rho_2\right)(1-\pi_2) \varepsilon(1) \\
+\kappa_{1, m}\left[A_{1, m}(1)-A_{1, m}(2)\right]\left[\left(\rho_1-\rho_2\right)(1-\pi_2)^2+\rho_2(1-\pi_2)\right]
\end{array}\right) x_t \\
&+\left(\begin{array}{c}
(\theta-1)\left[A_0(2)-A_0(1)\right]+\left[A_{0, m}(2)-A_{0, m}(1)\right] \\
+\left(-\frac{\theta}{\psi}+\theta-1\right)\left(\sigma_1-\sigma_2\right) \eta(1)(\pi_1+\pi_2-1)+\varphi_d\left(\sigma_1-\sigma_2\right) u(1)(\pi_1+\pi_2-1) \\
+(\theta-1) \kappa_1 A_1(2) \varphi_e\left(\sigma_1-\sigma_2\right) e(1)(\pi_1+\pi_2-1) \\
+(\theta-1) \kappa_1\left[A_0(1)-A_0(2)\right](\pi_1+\pi_2-1) \\
+2(\theta-1) \kappa_1\left[A_1(1)-A_1(2)\right] \varphi_e\left(\sigma_1-\sigma_2\right) e(1)(1-\pi_2)(\pi_1+\pi_2-1) \\
+(\theta-1) \kappa_1\left[A_1(1)-A_1(2)\right] \varphi_e \sigma_{\varepsilon, e}\left(\sigma_1-\sigma_2\right)(\pi_1+\pi_2-1) \\
+\kappa_{1, m} A_{1, m}(2) \varphi_e\left(\sigma_1-\sigma_2\right) e(1)(\pi_1+\pi_2-1) \\
+\kappa_{1, m}\left[A_{0, m}(1)-A_{0, m}(2)\right](\pi_1+\pi_2-1) \\
+2 \kappa_{1, m}\left[A_{1, m}(1)-A_{1, m}(2)\right] \varphi_e\left(\sigma_1-\sigma_2\right) e(1)(1-\pi_2)(\pi_1+\pi_2-1) \\
+\kappa_{1, m}\left[A_{1, m}(1)-A_{1, m}(2)\right] \varphi_e \sigma_{\varepsilon, e}\left(\sigma_1-\sigma_2\right)(\pi_1+\pi_2-1)
\end{array}\right) p_t \\
&+\left(\begin{array}{c}
(\theta-1)\left[A_1(2)-A_1(1)\right]+\left[A_{1, m}(2)-A_{1, m}(1)\right] \\
(\theta-1) \kappa_1 A_1(2)\left(\rho_1-\rho_2\right)(\pi_1+\pi_2-1) \\
+(\theta-1) \kappa_1\left[A_1(1)-A_1(2)\right]\left[2(1-\pi_2)(\pi_1+\pi_2-1)\left(\rho_1-\rho_2\right)+\rho_2(\pi_1+\pi_2-1)\right] \\
+(\theta-1) \kappa_1\left[A_1(1)-A_1(2)\right]\left(\rho_1-\rho_2\right) \varepsilon(1)(\pi_1+\pi_2-1) \\
\kappa_{1, m} A_{1, m}(2)\left(\rho_1-\rho_2\right)(\pi_1+\pi_2-1) \\
+\kappa_{1, m}\left[A_{1, m}(1)-A_{1, m}(2)\right]\left[2(1-\pi_2)(\pi_1+\pi_2-1)\left(\rho_1-\rho_2\right)+\rho_2(\pi_1+\pi_2-1)\right] \\
+\kappa_{1, m}\left[A_{1, m}(1)-A_{1, m}(2)\right]\left(\rho_1-\rho_2\right) \varepsilon(1)(\pi_1+\pi_2-1)
\end{array}\right) p_t x_t \\
&+(\pi_1+\pi_2-1)^2\left(\rho_1-\rho_2\right)\left\{(\theta-1) \kappa_1\left[A_1(1)-A_1(2)\right]+\kappa_{1, m}\left[A_{1, m}(1)-A_{1, m}(2)\right]\right\} p_t^2 x_t \\
&+\varphi_e\left(\sigma_1-\sigma_2\right) e(1)(\pi_1+\pi_2-1)^2\left\{(\theta-1) \kappa_1\left[A_1(1)-A_1(2)\right]+\kappa_{1, m}\left[A_{1, m}(1)-A_{1, m}(2)\right]\right\} p_t^2 \\
= & \;0
\end{aligned}
$$

As in Appendix A.2.1, we approximate the above expression to order x_t, p_t, and $p_t x_t$. Therefore, we expand the term p_t^2 as a Taylor series to first order around the unconditional mean, $\bar{p}$, of p_t.

Since the Euler equation holds for all observable states (x_t, p_t), we obtain the following 4 parameter restrictions:

Constant:

$$\left(\begin{array}{c}\theta\log\delta+(\theta-1)\kappa_0+\kappa_{0,m}-(\theta-1)A_0(2)-A_{0,m}(2)+\frac{1}{2}\Sigma' \\ \left(-\frac{\theta}{\psi}+\theta-1\right)(\mu+(\sigma_1-\sigma_2)\eta(1)(1-\pi_2))+\mu_d+\varphi_d(\sigma_1-\sigma_2)u(1)(1-\pi_2) \\ +(\theta-1)\kappa_1A_1(2)\varphi_e(\sigma_1-\sigma_2)e(1)(1-\pi_2)+(\theta-1)\kappa_1[A_0(1)-A_0(2)](1-\pi_2) \\ +(\theta-1)\kappa_1[A_1(1)-A_1(2)]\varphi_e(\sigma_1-\sigma_2)e(1)(1-\pi_2)^2+(\theta-1)\kappa_1A_0(2) \\ +(\theta-1)\kappa_1[A_1(1)-A_1(2)]\varphi_e\sigma_{\varepsilon,e}[(\sigma_1-\sigma_2)(1-\pi_2)+\sigma_2] \\ \kappa_{1,m}[A_{0,m}(1)-A_{0,m}(2)](1-\pi_2)+\kappa_{1,m}A_{0,m}(2) \\ +\kappa_{1,m}[A_{1,m}(1)-A_{1,m}(2)]\varphi_e\sigma_{\varepsilon,e}[(\sigma_1-\sigma_2)(1-\pi_2)+\sigma_2] \\ +\kappa_{1,m}[A_{1,m}(1)-A_{1,m}(2)]\varphi_e(\sigma_1-\sigma_2)e(1)(1-\pi_2)^2 \\ +\kappa_{1,m}A_{1,m}(1)\varphi_e(\sigma_1-\sigma_2)e(1)(1-\pi_2) \\ -\varphi_e(\sigma_1-\sigma_2)e(1)(\pi_1+\pi_2-1)^2\{(\theta-1)\kappa_1[A_1(1)-A_1(2)]+\kappa_{1,m}[A_{1,m}(1)-A_{1,m}(2)]\}\overline{p}^2\end{array}\right)=0$$

Coefficient of x_t:

$$\left(\begin{array}{c}-(\theta-1)A_1(2)-A_{1,m}(2)+\left(-\frac{\theta}{\psi}+\theta-1\right)+\phi \\ (\theta-1)\kappa_1A_1(2)[(\rho_1-\rho_2)(1-\pi_2)+\rho_2] \\ +(\theta-1)\kappa_1[A_1(1)-A_1(2)](\rho_1-\rho_2)(1-\pi_2)\varepsilon(1) \\ (\theta-1)\kappa_1[A_1(1)-A_1(2)][(\rho_1-\rho_2)(1-\pi_2)^2+\rho_2(1-\pi_2)] \\ +\kappa_{1,m}A_{1,m}(2)[(\rho_1-\rho_2)(1-\pi_2)+\rho_2] \\ +\kappa_{1,m}[A_{1,m}(1)-A_{1,m}(2)](\rho_1-\rho_2)(1-\pi_2)\varepsilon(1) \\ +\kappa_{1,m}[A_{1,m}(1)-A_{1,m}(2)][(\rho_1-\rho_2)(1-\pi_2)^2+\rho_2(1-\pi_2)]- \\ (\pi_1+\pi_2-1)^2(\rho_1-\rho_2)\{(\theta-1)\kappa_1[A_1(1)-A_1(2)]+\kappa_{1,m}[A_{1,m}(1)-A_{1,m}(2)]\}\overline{p}^2\end{array}\right)=0$$

Coefficient of p_t:

$$\left(\begin{array}{c}(\theta-1)[A_0(2)-A_0(1)]+[A_{0,m}(2)-A_{0,m}(1)] \\ +\left(-\frac{\theta}{\psi}+\theta-1\right)(\sigma_1-\sigma_2)\eta(1)(\pi_1+\pi_2-1)+\varphi_d(\sigma_1-\sigma_2)u(1)(\pi_1+\pi_2-1) \\ +(\theta-1)\kappa_1A_1(2)\varphi_e(\sigma_1-\sigma_2)e(1)(\pi_1+\pi_2-1) \\ +(\theta-1)\kappa_1[A_0(1)-A_0(2)](\pi_1+\pi_2-1) \\ +2(\theta-1)\kappa_1[A_1(1)-A_1(2)]\varphi_e(\sigma_1-\sigma_2)e(1)(1-\pi_2)(\pi_1+\pi_2-1) \\ +(\theta-1)\kappa_1[A_1(1)-A_1(2)]\varphi_e\sigma_{\varepsilon,e}(\sigma_1-\sigma_2)(\pi_1+\pi_2-1) \\ +\kappa_{1,m}A_{1,m}(2)\varphi_e(\sigma_1-\sigma_2)e(1)(\pi_1+\pi_2-1) \\ +\kappa_{1,m}[A_{0,m}(1)-A_{0,m}(2)](\pi_1+\pi_2-1) \\ +2\kappa_{1,m}[A_{1,m}(1)-A_{1,m}(2)]\varphi_e(\sigma_1-\sigma_2)e(1)(1-\pi_2)(\pi_1+\pi_2-1) \\ +\kappa_{1,m}[A_{1,m}(1)-A_{1,m}(2)]\varphi_e\sigma_{\varepsilon,e}(\sigma_1-\sigma_2)(\pi_1+\pi_2-1) \\ +2\varphi_e(\sigma_1-\sigma_2)e(1)(\pi_1+\pi_2-1)^2\{(\theta-1)\kappa_1[A_1(1)-A_1(2)]+\kappa_{1,m}[A_{1,m}(1)-A_{1,m}(2)]\}\overline{p}\end{array}\right)=0$$

Coefficient of $p_t x_t$:

$$\left(\begin{array}{c} (\theta-1)\left[A_1(2)-A_1(1)\right]+\left[A_{1,m}(2)-A_{1,m}(1)\right] \\ (\theta-1)\kappa_1 A_1(2)\left(\rho_1-\rho_2\right)\left(\pi_1+\pi_2-1\right) \\ +(\theta-1)\kappa_1\left[A_1(1)-A_1(2)\right]\left[2(1-\pi_2)(\pi_1+\pi_2-1)\left(\rho_1-\rho_2\right)+\rho_2(\pi_1+\pi_2-1)\right] \\ +(\theta-1)\kappa_1\left[A_1(1)-A_1(2)\right]\left(\rho_1-\rho_2\right)\varepsilon(1)(\pi_1+\pi_2-1) \\ \kappa_{1,m}A_{1,m}(2)\left(\rho_1-\rho_2\right)\left(\pi_1+\pi_2-1\right) \\ +\kappa_{1,m}\left[A_{1,m}(1)-A_{1,m}(2)\right]\left[2(1-\pi_2)(\pi_1+\pi_2-1)\left(\rho_1-\rho_2\right)+\rho_2(\pi_1+\pi_2-1)\right] \\ +\kappa_{1,m}\left[A_{1,m}(1)-A_{1,m}(2)\right]\left(\rho_1-\rho_2\right)\varepsilon(1)(\pi_1+\pi_2-1) \\ +2(\pi_1+\pi_2-1)^2\left(\rho_1-\rho_2\right)\left\{(\theta-1)\kappa_1\left[A_1(1)-A_1(2)\right]+\kappa_{1,m}\left[A_{1,m}(1)-A_{1,m}(2)\right]\right\}\overline{p} \end{array}\right)=0$$

The 4 linear equations can be solved to obtain the 4 parameters $A_{0,m}(1)$, $A_{0,m}(2)$, $A_{1,m}(1)$, and $A_{1,m}(2)$.

A.2.3 Riskfree Rate

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

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

Hence,

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

By Taylor series expansion up to quadratic terms, we obtain the following:

$$-r_{f,t}=\theta\log\delta+(\theta-1)\left(\kappa_0-z_t\right)+\left(-\frac{\theta}{\psi}+\theta-1\right)E\left[\Delta c_{t+1}|\digamma(t)\right]+(\theta-1)\kappa_1 E\left[z_{t+1}|\digamma(t)\right]+\frac{1}{2}\Sigma^{"},$$

where we approximate the conditional variance, $var\left(\left(-\frac{\theta}{\psi}+\theta-1\right)\Delta c_{t+1}+(\theta-1)\kappa_1 z_{t+1}|\digamma(t)\right)$, with the constant, $\Sigma^{"}$.

The above expression implies:

$$
\begin{aligned}
-r_{f,t} &= \theta \log \delta + (\theta-1)\kappa_0 - (\theta-1)\left\{p_t\left[A_0(1)+A_1(1)x_t\right]+(1-p_t)\left[A_0(2)+A_1(2)x_t\right]\right\} \\
&\quad + \left(-\frac{\theta}{\psi}+\theta-1\right)\left(\mu + x_t + (\sigma_1-\sigma_2)\,\eta(1) f(p_t)\right) \\
&\quad + (\theta-1)\kappa_1 \left(\begin{array}{c}
\left[A_0(1)-A_0(2)+(A_1(1)-A_1(2))\,\rho_2 x_t\right] f(p_t) \\
+A_0(2)+A_1(2)x_t\left(f(p_t)\rho_1+(1-f(p_t))\,\rho_2\right) \\
+\left(A_1(1)-A_1(2)\right)(\rho_1-\rho_2)\,\varepsilon(1) x_t f(p_t) \\
+\left(A_1(1)-A_1(2)\right)\varphi_e \sigma_{\varepsilon,e}\left\{\sigma_1 f(p_t)+\sigma_2\left(1-f(p_t)\right)\right\} \\
+\left(A_1(1)-A_1(2)\right)\left[(\rho_1-\rho_2)\,x_t+\varphi_e(\sigma_1-\sigma_2)\,e(1)\right]\left\{f(p_t)\right\}^2 \\
+\varphi_e(\sigma_1-\sigma_2)\,e(1)A_1(2)f(p_t)
\end{array}\right) \\
&\quad + \frac{1}{2}\Sigma^{"}
\end{aligned}
$$

Therefore, we obtain

$$
r_{f,t} = A_{0,f} + A_{1,f}x_t + A_{2,f}p_t + A_{3,f}x_t p_t
$$

where

$$
\begin{aligned}
-A_{0,f} &= \left(\begin{array}{c}
\theta\log\delta+(\theta-1)\kappa_0-(\theta-1)A_0(2)+\frac{1}{2}\Sigma^{"}+\left(-\frac{\theta}{\psi}+\theta-1\right)\left(\mu+(\sigma_1-\sigma_2)\,\eta(1)(1-\pi_2)\right) \\
+(\theta-1)\kappa_1 A_1(2)\varphi_e(\sigma_1-\sigma_2)\,e(1)(1-\pi_2)+(\theta-1)\kappa_1\left[A_0(1)-A_0(2)\right](1-\pi_2) \\
+(\theta-1)\kappa_1\left[A_1(1)-A_1(2)\right]\varphi_e(\sigma_1-\sigma_2)\,e(1)(1-\pi_2)^2+(\theta-1)\kappa_1 A_0(2) \\
+(\theta-1)\kappa_1\left[A_1(1)-A_1(2)\right]\varphi_e\sigma_{\varepsilon,e}\left[(\sigma_1-\sigma_2)(1-\pi_2)+\sigma_2\right] \\
-\varphi_e(\sigma_1-\sigma_2)\,e(1)(\pi_1+\pi_2-1)^2(\theta-1)\kappa_1\left[A_1(1)-A_1(2)\right]\overline{p}^2
\end{array}\right) \\
-A_{1,f} &= \left(\begin{array}{c}
-(\theta-1)A_1(2)+\left(-\frac{\theta}{\psi}+\theta-1\right) \\
(\theta-1)\kappa_1 A_1(2)\left[(\rho_1-\rho_2)(1-\pi_2)+\rho_2\right] \\
+(\theta-1)\kappa_1\left[A_1(1)-A_1(2)\right](\rho_1-\rho_2)(1-\pi_2)\varepsilon(1) \\
(\theta-1)\kappa_1\left[A_1(1)-A_1(2)\right]\left[(\rho_1-\rho_2)(1-\pi_2)^2+\rho_2(1-\pi_2)\right] \\
-(\pi_1+\pi_2-1)^2(\rho_1-\rho_2)(\theta-1)\kappa_1\left[A_1(1)-A_1(2)\right]\overline{p}^2
\end{array}\right) \\
-A_{2,f} &= \left(\begin{array}{c}
(\theta-1)\left[A_0(2)-A_0(1)\right]+\left(-\frac{\theta}{\psi}+\theta-1\right)(\sigma_1-\sigma_2)\,\eta(1)(\pi_1+\pi_2-1) \\
+(\theta-1)\kappa_1 A_1(2)\varphi_e(\sigma_1-\sigma_2)\,e(1)(\pi_1+\pi_2-1) \\
+(\theta-1)\kappa_1\left[A_0(1)-A_0(2)\right](\pi_1+\pi_2-1) \\
+2(\theta-1)\kappa_1\left[A_1(1)-A_1(2)\right]\varphi_e(\sigma_1-\sigma_2)\,e(1)(1-\pi_2)(\pi_1+\pi_2-1) \\
+(\theta-1)\kappa_1\left[A_1(1)-A_1(2)\right]\varphi_e\sigma_{\varepsilon,e}(\sigma_1-\sigma_2)(\pi_1+\pi_2-1) \\
+2\varphi_e(\sigma_1-\sigma_2)\,e(1)(\pi_1+\pi_2-1)^2(\theta-1)\kappa_1\left[A_1(1)-A_1(2)\right]\overline{p}
\end{array}\right) \\
-A_{3,f} &= \left(\begin{array}{c}
(\theta-1)\left[A_1(2)-A_1(1)\right] \\
(\theta-1)\kappa_1 A_1(2)(\rho_1-\rho_2)(\pi_1+\pi_2-1) \\
+(\theta-1)\kappa_1\left[A_1(1)-A_1(2)\right]\left[2(1-\pi_2)(\pi_1+\pi_2-1)(\rho_1-\rho_2)+\rho_2(\pi_1+\pi_2-1)\right] \\
+(\theta-1)\kappa_1\left[A_1(1)-A_1(2)\right](\rho_1-\rho_2)\,\varepsilon(1)(\pi_1+\pi_2-1) \\
+2(\pi_1+\pi_2-1)^2(\rho_1-\rho_2)(\theta-1)\kappa_1\left[A_1(1)-A_1(2)\right]\overline{p}
\end{array}\right)
\end{aligned}
$$

A.2.4 Equity Premium

Using the log-linearized return on the market portfolio in equation (16) and noting that the log price-dividend ratio of the market is given by equation (18), we have

$$\begin{aligned} E\left(r_{m,t+1}|\digamma(t)\right) &= \kappa_{0,m}+\kappa_{1,m}E\left[z_{m,t+1}|\digamma(t)\right]-z_{m,t}+E\left[\Delta d_{t+1}|\digamma(t)\right] \\ &= B_0+B_1x_t+B_2p_t+B_3p_tx_t \end{aligned}$$

where

$$\begin{aligned}
B_0 &= \left(\begin{array}{c} \kappa_{0,m}-A_{0,m}(2)+\mu_d+\varphi_d\left(\sigma_1-\sigma_2\right)u(1)(1-\pi_2) \\ +\kappa_{1,m}A_{1,m}(2)\varphi_e\left(\sigma_1-\sigma_2\right)e(1)(1-\pi_2)+\kappa_{1,m}\left[A_{0,m}(1)-A_{0,m}(2)\right](1-\pi_2) \\ +\kappa_{1,m}\left[A_{1,m}(1)-A_{1,m}(2)\right]\varphi_e\left(\sigma_1-\sigma_2\right)e(1)(1-\pi_2)^2+\kappa_{1,m}A_{0,m}(2) \\ +\kappa_{1,m}\left[A_{1,m}(1)-A_{1,m}(2)\right]\varphi_e\sigma_{\varepsilon,e}\left[\left(\sigma_1-\sigma_2\right)(1-\pi_2)+\sigma_2\right] \\ -\varphi_e\left(\sigma_1-\sigma_2\right)e(1)(\pi_1+\pi_2-1)^2\kappa_{1,m}\left[A_{1,m}(1)-A_{1,m}(2)\right]\overline{p}^2 \end{array}\right) \\
B_1 &= \left(\begin{array}{c} -A_{1,m}(2)+\phi+\kappa_{1,m}A_{1,m}(2)\left[\left(\rho_1-\rho_2\right)(1-\pi_2)+\rho_2\right] \\ +\kappa_{1,m}\left[A_{1,m}(1)-A_{1,m}(2)\right]\left(\rho_1-\rho_2\right)(1-\pi_2)\varepsilon(1) \\ +\kappa_{1,m}\left[A_{1,m}(1)-A_{1,m}(2)\right]\left[\left(\rho_1-\rho_2\right)(1-\pi_2)^2+\rho_2(1-\pi_2)\right] \\ -(\pi_1+\pi_2-1)^2\left(\rho_1-\rho_2\right)\kappa_{1,m}\left[A_{1,m}(1)-A_{1,m}(2)\right]\overline{p}^2 \end{array}\right) \\
B_2 &= \left(\begin{array}{c} \left[A_{0,m}(2)-A_{0,m}(1)\right]+\varphi_d\left(\sigma_1-\sigma_2\right)u(1)(\pi_1+\pi_2-1) \\ +\kappa_{1,m}A_{1,m}(2)\varphi_e\left(\sigma_1-\sigma_2\right)e(1)(\pi_1+\pi_2-1) \\ +\kappa_{1,m}\left[A_{0,m}(1)-A_{0,m}(2)\right](\pi_1+\pi_2-1) \\ +2\kappa_{1,m}\left[A_{1,m}(1)-A_{1,m}(2)\right]\varphi_e\left(\sigma_1-\sigma_2\right)e(1)(1-\pi_2)(\pi_1+\pi_2-1) \\ +\kappa_{1,m}\left[A_{1,m}(1)-A_{1,m}(2)\right]\varphi_e\sigma_{\varepsilon,e}\left(\sigma_1-\sigma_2\right)(\pi_1+\pi_2-1) \\ +2\varphi_e\left(\sigma_1-\sigma_2\right)e(1)(\pi_1+\pi_2-1)^2\kappa_{1,m}\left[A_{1,m}(1)-A_{1,m}(2)\right]\overline{p} \end{array}\right) \\
B_3 &= \left(\begin{array}{c} \left[A_{1,m}(2)-A_{1,m}(1)\right] \\ \kappa_{1,m}A_{1,m}(2)\left(\rho_1-\rho_2\right)(\pi_1+\pi_2-1) \\ +\kappa_{1,m}\left[A_{1,m}(1)-A_{1,m}(2)\right]\left[2(1-\pi_2)(\pi_1+\pi_2-1)\left(\rho_1-\rho_2\right)+\rho_2(\pi_1+\pi_2-1)\right] \\ +\kappa_{1,m}\left[A_{1,m}(1)-A_{1,m}(2)\right]\left(\rho_1-\rho_2\right)\varepsilon(1)(\pi_1+\pi_2-1) \\ +2(\pi_1+\pi_2-1)^2\left(\rho_1-\rho_2\right)\kappa_{1,m}\left[A_{1,m}(1)-A_{1,m}(2)\right]\overline{p} \end{array}\right)
\end{aligned}$$

Now, the risk free rate is given by equation (19)

$$r_{f,t}=A_{0,f}+A_{1,f}x_t+A_{2,f}p_t+A_{3,f}x_tp_t.$$

Hence, the equity premium is given by

$$E\left[\left(r_{m,t+1}-r_{f,t}\right)|\digamma(t)\right]=E_0+E_1x_t+E_2p_t+E_3p_tx_t,$$

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

A.2.5 Predictive Implications for Consumption and Dividend Growth

Equation (2) implies that the expected consumption growth rate is given by

$$
\begin{aligned}
& E\left(\Delta c_{t+1}|F(t)\right) \\
= \; & \mu + x_t + E\left[\sigma_{s_{t+1}}\eta_{t+1}|F(t)\right] \\
= \; & \mu + x_t + (\sigma_1 - \sigma_2)\,\eta(1) f(p_t) \\
= \; & \mu + (\sigma_1 - \sigma_2)\,\eta(1)\,(1-\pi_2) + x_t + (\sigma_1 - \sigma_2)\,\eta(1)\,(\pi_1 + \pi_2 - 1)\,p_t.
\end{aligned}
$$

Similarly, the expected dividend growth rate is given by

$$
\begin{aligned}
& E\left(\Delta d_{t+1}|F(t)\right) \\
= \; & \mu_d + \phi x_t + \varphi_d E\left[\sigma_{s_{t+1}} u_{t+1}|F(t)\right] \\
= \; & \mu_d + \phi x_t + \varphi_d\,(\sigma_1 - \sigma_2)\,u(1) f(p_t) \\
= \; & \mu_d + \varphi_d\,(\sigma_1 - \sigma_2)\,u(1)\,(1-\pi_2) + \phi x_t + \varphi_d\,(\sigma_1 - \sigma_2)\,u(1)\,(\pi_1 + \pi_2 - 1)\,p_t.
\end{aligned}
$$

Therefore, the expected consumption and dividend growth rates are both linear functions of the state variables, x_t and p_t.

Finally, the model implies that the conditional variance of the aggregate consumption and dividend growth rates are functions of the probability, p_t, alone:

$$
\begin{aligned}
Var\left(\Delta c_{t+1}|F(t)\right) &= Var\left[\sigma_{s_{t+1}}\eta_{t+1}|F(t)\right] \\
&= E\left[\sigma_{s_{t+1}}^2\eta_{t+1}^2|F(t)\right] - \left(E\left[\sigma_{s_{t+1}}\eta_{t+1}|F(t)\right]\right)^2 \\
&= c_1 + d_1 p_t + e_1 p_t^2,
\end{aligned}
$$

where

$$
\begin{aligned}
c_1 &= \sigma_2^2 + \left(\sigma_1^2 - \sigma_2^2\right)(1-\pi_2) - (\sigma_1 - \sigma_2)^2\,\eta(1)^2\,(1-\pi_2)^2, \\
d_1 &= \left[\left(\sigma_1^2 - \sigma_2^2\right)(\pi_1 + \pi_2 - 1) - 2\,(\sigma_1 - \sigma_2)^2\,\eta(1)^2\,(1-\pi_2)\,(\pi_1 + \pi_2 - 1)\right], \\
e_1 &= -\,(\sigma_1 - \sigma_2)^2\,\eta(1)^2\,(\pi_1 + \pi_2 - 1)^2.
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
Var\left(\Delta d_{t+1}|F(t)\right) &= Var\left[\varphi_d\sigma_{s_{t+1}} u_{t+1}|F(t)\right] \\
&= \varphi_d^2 E\left[\sigma_{s_{t+1}}^2 u_{t+1}^2|F(t)\right] - \varphi_d^2\left(E\left[\sigma_{s_{t+1}} u_{t+1}|F(t)\right]\right)^2 \\
&= c_2 + d_2 p_t + e_2 p_t^2,
\end{aligned}
$$

where

$$
\begin{aligned}
c_2 &= \varphi_d^2\left[\sigma_2^2 + \left(\sigma_1^2 - \sigma_2^2\right)(1-\pi_2) - (\sigma_1 - \sigma_2)^2\,u(1)^2\,(1-\pi_2)^2\right], \\
d_2 &= \varphi_d^2\left[\left(\sigma_1^2 - \sigma_2^2\right)(\pi_1 + \pi_2 - 1) - 2\,(\sigma_1 - \sigma_2)^2\,u(1)^2\,(1-\pi_2)\,(\pi_1 + \pi_2 - 1)\right], \\
e_2 &= -\varphi_d^2\,(\sigma_1 - \sigma_2)^2\,u(1)^2\,(\pi_1 + \pi_2 - 1)^2.
\end{aligned}
$$

Expanding the term p_t^2 as a Taylor series to first order around the unconditional mean, $\overline{p}$, of p_t, we obtain equations (24) and (25) for the conditional variance of the aggregate consumption and dividend growth rates, respectively.

A.2.6 Pricing Kernel

The pricing kernel is given by equation (14),

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

Now, the log-linearization in equation (15),

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

along with the solution for z_t,

$$z_t = p_t \left[A_0(1) + A_1(1) x_t \right] + (1 - p_t) \left[A_0(2) + A_1(2) x_t \right],$$

together imply that,

$$\begin{aligned} m_{t+1} &= \theta \log \delta - \frac{\theta}{\psi} \Delta c_{t+1} + (\theta - 1)\kappa_0 \\ &\quad + (\theta - 1)\kappa_1 p_{t+1} \left[A_0(1) + A_1(1) x_{t+1} \right] \\ &\quad + (\theta - 1)\kappa_1 (1 - p_{t+1}) \left[A_0(2) + A_1(2) x_{t+1} \right] \\ &\quad - (\theta - 1) p_t \left[A_0(1) + A_1(1) x_t \right] - (\theta - 1)(1 - p_t) \left[A_0(2) + A_1(2) x_t \right] \\ &\quad + (\theta - 1)\Delta c_{t+1}. \end{aligned}$$

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,

$$\begin{aligned} c_0 &= \theta \log(\delta) + (\theta - 1)\kappa_0 + (\theta - 1)(\kappa_1 - 1) A_0(2), \\ c_1 &= -\frac{\theta}{\psi} + \theta - 1, \\ c_2 &= (\theta - 1)\kappa_1 \left[A_0(1) - A_0(2) \right], \\ c_3 &= -(\theta - 1) \left[A_0(1) - A_0(2) \right], \\ c_4 &= (\theta - 1)\kappa_1 A_1(2), \\ c_5 &= -(\theta - 1) A_1(2), \\ c_6 &= (\theta - 1)\kappa_1 \left[A_1(1) - A_1(2) \right], \\ c_7 &= -(\theta - 1) \left[A_1(1) - A_1(2) \right]. \end{aligned}$$

This expression for the pricing kernel involves the state variables, x_t and p_t. These are latent to the econometrician. However, note that the log price-dividend ratio of the aggregate stock market, $z_{m,t}$, and the risk free rate, $r_{f,t}$, are functions only of these two latent state variables (equations (18) and (19)),

$$
\begin{aligned}
z_{m,t} &= p_t \left[A_{0,m}(1) + A_{1,m}(1)x_t\right] + (1-p_t)\left[A_{0,m}(2) + A_{1,m}(2)x_t\right] \\
&= A_{0,m}(2) + A_{1,m}(2)x_t + \left[A_{0,m}(1) - A_{0,m}(2)\right] p_t + \left[A_{1,m}(1) - A_{1,m}(2)\right] p_t x_t,
\end{aligned}
$$

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

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, (18) implies

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

Substituting (44) into (19), and simplifying gives the following quadratic equation for p_t:

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

where

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

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

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

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.3 Time-Series Moments

We compute the unconditional moments of the aggregate consumption and dividend growth rates. To do that, we first compute the unconditional expectations of the state variables x_t and p_t, and their cross-product $x_t p_t$. Note that

$$
\begin{aligned}
E\left[x_{t+1}|\digamma(t)\right] &= E\left[\rho_{s_{t+1}}|\digamma(t)\right]x_t + \varphi_e E\left[\sigma_{s_{t+1}}e_{t+1}|\digamma(t)\right] \\
&= \left[f(p_t)\rho_1 + (1-f(p_t))\,\rho_2\right]x_t + \varphi_e\left(\sigma_1-\sigma_2\right)e(1)f(p_t) \\
&= \varphi_e\left(\sigma_1-\sigma_2\right)e(1)(1-\pi_2) + \left[(1-\pi_2)\rho_1+\pi_2\rho_2\right]x_t \\
&\quad +\varphi_e\left(\sigma_1-\sigma_2\right)e(1)(\pi_1+\pi_2-1)p_t + (\pi_1+\pi_2-1)(\rho_1-\rho_2)p_t x_t
\end{aligned}
$$

Taking expectations of the two sides of the above equation gives

$$
\begin{aligned}
& E\left(x_t\right)\left[1-(1-\pi_2)\rho_1-\pi_2\rho_2\right] \\
= \;& \varphi_e\left(\sigma_1-\sigma_2\right)e(1)(1-\pi_2) + \varphi_e\left(\sigma_1-\sigma_2\right)e(1)(\pi_1+\pi_2-1)E\left(p_t\right) \\
& +(\pi_1+\pi_2-1)(\rho_1-\rho_2)E\left(p_t x_t\right) \qquad (47)
\end{aligned}
$$

Also,

$$
\begin{aligned}
E\left[x_{t+1}p_{t+1}|\digamma(t)\right] &= E\left[\left(\rho_{s_{t+1}}x_t+\varphi_e\sigma_{s_{t+1}}e_{t+1}\right)\left((1-\pi_2)+(\pi_1+\pi_2-1)p_t+\varepsilon_{t+1}\right)|\digamma(t)\right] \\
&= (1-\pi_2)\left[(1-\pi_2)\rho_1+\pi_2\rho_2\right]x_t + (1-\pi_2)(\pi_1+\pi_2-1)(\rho_1-\rho_2)p_t x_t \\
&\quad +\varphi_e\left(\sigma_1-\sigma_2\right)e(1)(1-\pi_2)^2 + (1-\pi_2)\varphi_e\left(\sigma_1-\sigma_2\right)e(1)(\pi_1+\pi_2-1)p_t \\
&\quad +(\pi_1+\pi_2-1)\left[(1-\pi_2)\rho_1+\pi_2\rho_2\right]p_t x_t + (\pi_1+\pi_2-1)^2(\rho_1-\rho_2)p_t^2 x_t \\
&\quad +(\pi_1+\pi_2-1)\varphi_e\left(\sigma_1-\sigma_2\right)e(1)(1-\pi_2)p_t + \varphi_e\left(\sigma_1-\sigma_2\right)e(1)(\pi_1+\pi_2-1)^2 p_t^2 \\
&\quad +(\rho_1-\rho_2)\varepsilon(1)(1-\pi_2)x_t + (\rho_1-\rho_2)\varepsilon(1)(\pi_1+\pi_2-1)p_t x_t \\
&\quad +\varphi_e\sigma_{\varepsilon e}\sigma_2 + \varphi_e\sigma_{\varepsilon e}\left(\sigma_1-\sigma_2\right)\left(1-\pi_2\right) + \varphi_e\sigma_{\varepsilon e}\left(\sigma_1-\sigma_2\right)\left(\pi_1+\pi_2-1\right)p_t
\end{aligned}
$$

Taking a first order Taylor series approximation of $p_t^2 \approx \left[E(p_t)\right]^2 + 2E(p_t)(p_t - E(p_t))$, and then taking expectations of both sides of the above expression gives

$$
\begin{aligned}
E\left(x_{t+1}p_{t+1}\right) &= \begin{pmatrix} \varphi_e\left(\sigma_1-\sigma_2\right)e(1)(1-\pi_2)^2 - \varphi_e\left(\sigma_1-\sigma_2\right)e(1)(\pi_1+\pi_2-1)^2\left[E(p_t)\right]^2 \\ +\varphi_e\sigma_{\varepsilon e}\sigma_2 + \varphi_e\sigma_{\varepsilon e}\left(\sigma_1-\sigma_2\right)\left(1-\pi_2\right) \end{pmatrix} \\
&\quad + \begin{pmatrix} (1-\pi_2)\left[(1-\pi_2)\rho_1+\pi_2\rho_2\right] - (\pi_1+\pi_2-1)^2(\rho_1-\rho_2)\left[E(p_t)\right]^2 \\ +(\rho_1-\rho_2)\varepsilon(1)(1-\pi_2) \end{pmatrix} E(x_t) \\
&\quad + \begin{pmatrix} 2(1-\pi_2)\varphi_e\left(\sigma_1-\sigma_2\right)e(1)(\pi_1+\pi_2-1) + \varphi_e\sigma_{\varepsilon e}\left(\sigma_1-\sigma_2\right)\left(\pi_1+\pi_2-1\right) \\ +2(\pi_1+\pi_2-1)^2\varphi_e\left(\sigma_1-\sigma_2\right)e(1)E(p_t) \end{pmatrix} E(p_t) \\
&\quad + \begin{pmatrix} (1-\pi_2)(\pi_1+\pi_2-1)(\rho_1-\rho_2) + (\pi_1+\pi_2-1)\left[(1-\pi_2)\rho_1+\pi_2\rho_2\right] \\ +2(\pi_1+\pi_2-1)^2(\rho_1-\rho_2)E(p_t) + (\rho_1-\rho_2)\varepsilon(1)(\pi_1+\pi_2-1) \end{pmatrix} E(p_t x_t) \qquad (48)
\end{aligned}
$$

Note that $E(p_t) = \frac{1-\pi_2}{2-\pi_1-\pi_2}$. Therefore, the equations (47) and (48) can be solved to obtain $E(x_t)$ and $E(p_t x_t)$.

Now,

$$
E\left(\Delta c_{t+1}\right) = \mu + E(x_t) + E\left[\sigma_{s_{t+1}}\eta_{t+1}\right] \qquad (49)
$$

Therefore, the unconditional mean of the consumption growth rate can be computed using the expression for $E(x_t)$ obtained above and the expression for $E\left[\sigma_{s_{t+1}}\eta_{t+1}|\digamma(t)\right] = \left(\sigma_1 - \sigma_2\right)\eta(1)f(p_t)$ obtained in Appendix A.2 (implying that $E\left[\sigma_{s_{t+1}}\eta_{t+1}\right] = \left(\sigma_1 - \sigma_2\right)\eta(1)(1-\pi_2) + \left(\sigma_1 - \sigma_2\right)\eta(1)(\pi_1 + \pi_2 - 1)E(p_t))$.

Similarly, we obtain

$$E\left(\Delta d_{t+1}\right) = \mu_d + \phi E(x_t) + \varphi_d E\left[\sigma_{s_{t+1}}u_{t+1}\right] \tag{50}$$

by noting from Appendix A.2 that $E\left[\sigma_{s_{t+1}}u_{t+1}|\digamma(t)\right] = \left(\sigma_1 - \sigma_2\right)u(1)f(p_t)$ implying that $E\left[\sigma_{s_{t+1}}u_{t+1}\right] = \left(\sigma_1 - \sigma_2\right)u(1)(1-\pi_2) + \left(\sigma_1 - \sigma_2\right)u(1)(\pi_1 + \pi_2 - 1)E(p_t)$.

Next, we compute the unconditional variances of consumption and dividend growth rates. Note that,

$$Var\left(\Delta c_{t+1}\right) = Var(x_t) + Var\left(\sigma_{s_{t+1}}\eta_{t+1}\right) + 2Cov\left(x_t, \sigma_{s_{t+1}}\eta_{t+1}\right). \tag{51}$$

Consider first the second term of equation (51):

$$\begin{aligned}
Var\left(\sigma_{s_{t+1}}\eta_{t+1}\right) &= E\left(\sigma^2_{s_{t+1}}\eta^2_{t+1}\right) - \left\{E\left(\sigma_{s_{t+1}}\eta_{t+1}\right)\right\}^2 \\
&= E\left[E\left(\sigma^2_{s_{t+1}}\eta^2_{t+1}|s_{t+1}\right)\right] - \left\{E\left(\sigma_{s_{t+1}}\eta_{t+1}\right)\right\}^2 \\
&= \frac{1-\pi_2}{2-\pi_1-\pi_2}\sigma_1^2 + \frac{1-\pi_1}{2-\pi_1-\pi_2}\sigma_2^2 \\
&\quad - \left\{\left(\sigma_1 - \sigma_2\right)\eta(1)(1-\pi_2) + \left(\sigma_1 - \sigma_2\right)\eta(1)(\pi_1+\pi_2-1)E(p_t)\right\}^2
\end{aligned}$$

Consider next the third term of equation (51):

$$\begin{aligned}
Cov\left(x_t, \sigma_{s_{t+1}}\eta_{t+1}\right) &= E\left(x_t\sigma_{s_{t+1}}\eta_{t+1}\right) - E(x_t)E(\sigma_{s_{t+1}}\eta_{t+1}) \\
&= E\left[E(x_t\sigma_{s_{t+1}}\eta_{t+1}|\digamma(t))\right] - E(x_t)E(\sigma_{s_{t+1}}\eta_{t+1}) \\
&= E\left[x_t\left(\sigma_1 - \sigma_2\right)\eta(1)f(p_t)\right] - E(x_t)E(\sigma_{s_{t+1}}\eta_{t+1}) \\
&= \left(\sigma_1 - \sigma_2\right)\eta(1)(\pi_1+\pi_2-1)\left[E(p_tx_t) - E(p_t)E(x_t)\right]
\end{aligned}$$

Finally, consider the first term of equation (51):

$$\begin{aligned}
Var(x_{t+1}) &\equiv E(x_t^2) - [E(x_t)]^2 \\
&= Var(\rho_{s_{t+1}}x_t) + Var\left(\varphi_e\sigma_{s_{t+1}}e_{t+1}\right) + 2Cov\left(\rho_{s_{t+1}}x_t, \varphi_e\sigma_{s_{t+1}}e_{t+1}\right)
\end{aligned} \tag{52}$$

Now,

$$\begin{aligned}
Var\left(\varphi_e\sigma_{s_{t+1}}e_{t+1}\right) &= \varphi_e^2\left(\frac{1-\pi_2}{2-\pi_1-\pi_2}\sigma_1^2 + \frac{1-\pi_1}{2-\pi_1-\pi_2}\sigma_2^2\right) \\
&\quad -\varphi_e^2\left\{\left(\sigma_1 - \sigma_2\right)e(1)(1-\pi_2) + \left(\sigma_1 - \sigma_2\right)e(1)(\pi_1+\pi_2-1)E(p_t)\right\}^2
\end{aligned}$$

$$
\begin{aligned}
Cov\left(\rho_{s_{t+1}}x_t, \varphi_e\sigma_{s_{t+1}}e_{t+1}\right) &= \varphi_e E\left(\rho_{s_{t+1}}\sigma_{s_{t+1}}e_{t+1}x_t\right) - E\left(\rho_{s_{t+1}}x_t\right)E\left(\varphi_e\sigma_{s_{t+1}}e_{t+1}\right) \\
&= \varphi_e E\left[x_t E\left(\rho_{s_{t+1}}\sigma_{s_{t+1}}e_{t+1}|F(t)\right)\right] - E\left(\rho_{s_{t+1}}x_t\right)E\left(\varphi_e\sigma_{s_{t+1}}e_{t+1}\right) \\
&= \varphi_e E\left[(\sigma_1\rho_1 - \sigma_2\rho_2)\, e(1) f(p_t) x_t\right] \\
&\quad -\varphi_e E\left[\left(f(p_t)\rho_1 - (1-f(p_t))\rho_2\right) x_t\right](\sigma_1 - \sigma_2)\, e(1) E\left[f(p_t)\right] \\
&= \varphi_e\left(\sigma_1\rho_1 - \sigma_2\rho_2\right) e(1)(1-\pi_1)E(x_t) \\
&\quad +\varphi_e\left(\sigma_1\rho_1 - \sigma_2\rho_2\right) e(1)(\pi_1+\pi_2-1)E(p_t x_t) \\
&\quad -\varphi_e\left(\sigma_1 - \sigma_2\right) e(1) E\left[\left(f(p_t)\rho_1 - (1-f(p_t))\rho_2\right) x_t\right] E\left[f(p_t)\right]
\end{aligned}
$$

$$
\begin{aligned}
Var(\rho_{s_{t+1}}x_t) &= E(\rho^2_{s_{t+1}}x_t^2) - \left[E(\rho_{s_{t+1}}x_t)\right]^2 \\
&= E\left[\left(f(p_t)\rho_1^2 + (1-f(p_t))\rho_2^2\right)x_t^2\right] - \left\{E\left[\left(f(p_t)\rho_1 + (1-f(p_t))\rho_2\right)x_t\right]\right\}^2 \\
&= E\left[\left(\rho_1^2 - \rho_2^2\right)(1-\pi_2)x_t^2 + \left(\rho_1^2 - \rho_2^2\right)(\pi_1+\pi_2-1)p_t x_t^2 + \rho_2^2 x_t^2\right] \\
&\quad - \left\{E\left[\left(\rho_1 - \rho_2\right)(1-\pi_2)x_t + \left(\rho_1 - \rho_2\right)(\pi_1+\pi_2-1)p_t x_t + \rho_2 x_t\right]\right\}^2
\end{aligned}
$$

Approximating $E(p_t x_t^2) \approx E(p_t)E(x_t^2)$, we solve equation (52) for $E(x_t^2)$. Substituting the expressions for $Var(x_t)$, $Var\left(\sigma_{s_{t+1}}\eta_{t+1}\right)$, and $Cov\left(x_t, \sigma_{s_{t+1}}\eta_{t+1}\right)$ into equation (51) gives the unconditional variance of consumption growth. Similarly, the unconditional variance of the dividend growth rate may be obtained as:

$$
Var\left(\Delta d_{t+1}\right) = \phi^2 Var(x_t) + \varphi_d^2 Var\left(\sigma_{s_{t+1}}u_{t+1}\right) + 2\phi\varphi_d Cov\left(x_t, \sigma_{s_{t+1}}u_{t+1}\right). \tag{53}
$$

Finally, we have

$$
\begin{aligned}
Cov\left(\Delta c_{t+1}, \Delta d_{t+1}\right) &= Cov\left(x_t + \sigma_{s_{t+1}}\eta_{t+1}, \phi x_t + \varphi_d\sigma_{s_{t+1}}u_{t+1}\right) \\
&= \phi Var(x_t) + Cov\left(x_t, \varphi_d\sigma_{s_{t+1}}u_{t+1}\right) + Cov\left(\sigma_{s_{t+1}}\eta_{t+1}, \phi x_t\right) \\
&\quad + Cov\left(\sigma_{s_{t+1}}\eta_{t+1}, \varphi_d\sigma_{s_{t+1}}u_{t+1}\right)
\end{aligned} \tag{54}
$$

where $Cov\left(x_t, \varphi_d\sigma_{s_{t+1}}u_{t+1}\right) = \varphi_d\left\{(\sigma_1 - \sigma_2)\, u(1)(\pi_1+\pi_2-1)\left[E(p_t x_t) - E(p_t)E(x_t)\right]\right\}$,
$Cov\left(x_t, \phi\sigma_{s_{t+1}}\eta_{t+1}\right) = \phi\left\{(\sigma_1 - \sigma_2)\,\eta(1)(\pi_1+\pi_2-1)\left[E(p_t x_t) - E(p_t)E(x_t)\right]\right\}$, and
$Cov\left(\sigma_{s_{t+1}}\eta_{t+1}, \varphi_d\sigma_{s_{t+1}}u_{t+1}\right) = \varphi_d p\eta(1)u(1)\left[\sigma_1^2 - \sigma_2^2\frac{p}{1-p}\right] - \varphi_d E\left(\sigma_{s_{t+1}}\eta_{t+1}\right)E\left(\sigma_{s_{t+1}}u_{t+1}\right)$.

A.4 Simulation Design

We assume that the error term, ε_t, has the following Bernoulli distribution, conditional on the economy being in the first regime at date t:

$$(\varepsilon_t|s_t=1)=\begin{cases}\max\left(-\left(1-\pi_2\right),-\frac{(1-\pi_1)^2}{1-\pi_2}\right), & \text{with prob}=\frac{\varepsilon(1)-(1-\pi_1)}{\max\left(-(1-\pi_2),-\frac{(1-\pi_1)^2}{1-\pi_2}\right)-(1-\pi_1)}\\ 1-\pi_1, & \text{with prob}=1-\frac{\varepsilon(1)-(1-\pi_1)}{\max\left(-(1-\pi_2),-\frac{(1-\pi_1)^2}{1-\pi_2}\right)-(1-\pi_1)}\end{cases},$$

and that the error term has the following Bernoulli distribution, conditional on the economy being in the second regime at date t:

$$(\varepsilon_t|s_t=2)=\begin{cases}\max\left(-\left(1-\pi_2\right),-\frac{(1-\pi_1)^2}{1-\pi_2}\right), & \text{with prob}=\frac{-\varepsilon(1)\frac{1-\pi_2}{1-\pi_1}-(1-\pi_1)}{\max\left(-(1-\pi_2),-\frac{(1-\pi_1)^2}{1-\pi_2}\right)-(1-\pi_1)}\\ 1-\pi_1, & \text{with prob}=1-\frac{-\varepsilon(1)\frac{1-\pi_2}{1-\pi_1}-(1-\pi_1)}{\max\left(-(1-\pi_2),-\frac{(1-\pi_1)^2}{1-\pi_2}\right)-(1-\pi_1)}\end{cases}.$$

Note that, in the model, we do not take a stand on the content of the information set, $\mathcal{F}(t)$, that the consumer uses to form his belief, p_t. In other words, the distribution of the error term, ε_t, of the probability evolution equation (7) is left unspecified. The assumption of a Bernoulli distribution in the simulations is just one choice among a set of many possible specifications of the distribution.

We further assume that, conditional on the economy being in the first regime at date t, the distribution of each of the error terms $\{e_t,\ \eta_t,\ u_t\}$ is independent of each other and is normal:

$$(y_t|s_t=1)\sim N\left(y\left(1\right),1\right),\quad y=e,\eta,u,$$

and that, conditional on the economy being in the second regime at date t, the distribution of each of the error terms $\{e_t,\ \eta_t,\ u_t\}$ is independent of each other and is normal:

$$(y_t|s_t=2)\sim N\left(-y\left(1\right)\frac{1-\pi_2}{1-\pi_1},1\right),\quad y=e,\eta,u.$$

We generate each history as follows: (i) we draw from the Markov process in equation (5) and generate a time series of the regime; (ii) conditional on the time series of the regime, we generate time series of the state variables, x_t and p_t, using equations (1) and (7), respectively; (iii) conditional on the time series of the state variables, we generate time series of aggregate consumption and dividend growth rates, using equations (2) and (3), respectively; and (iv) we generate the time series of the price-dividend ratio, risk free rate, and market return, using equations (18), (19), and (20), respectively. We repeat this procedure $10,000$ times and generate $10,000$ histories.

We compute the mean and volatility of the aggregate consumption and dividend growth rates and the risk free rate, market-wide price-dividend ratio, market return, and equity premium in each history. We also split each history into two where the first subsample corresponds to those time periods when $p_t > 0.5$ while the second subsample corresponds to those time periods when $p_t < 0.5$. We compute the mean and volatility of the aggregate consumption and dividend growth rates and the risk free rate, market-wide price-dividend ratio, market return, and equity premium in each subsample.

Table 1: Parameter Estimates, 1930-2009

μ	μ_d	ϕ	φ_d	ρ_1	ρ_2	φ_e	σ_1	σ_2	π_1	π_2
0.005 (0.101)	-0.044 (0.330)	3.5 (0.15)	2.0 (0.26)	0.94 (0.46)	0.6 (0.90)	0.46 (0.90)	0.005 (0.073)	0.035 (0.080)	0.95 (0.23)	0.85 (0.63)
$\eta(1)$	$e(1)$	$u(1)$	$\varepsilon(1)$	$\sigma_{\varepsilon e}$						
0 (0.72)	-0.25 (0.20)	0 (0.39)	0.02 (0.17)	-0.5 (0.39)						
δ	γ	ψ								
0.976 (0.342)	12 (20.2)	0.9 (0.044)								

	Data	*Model*		*Data*	*Model*
$E(r_f)$	0.008 (0.008)	0.012 [−.008,0.030]	$E(\Delta c)$	0.015 (0.003)	0.015 [−0.002,0.032]
$\sigma(r_f)$	0.050 (0.008)	0.018 [0.007,0.028]	$sd(\Delta c)$	0.025 (0.004)	0.024 [0.010,0.035]
$E(r_m - r_f)$	0.058 (0.020)	0.033 [−.000,0.066]	$E(\Delta d)$	0.017 (0.013)	0.008 [−0.071,0.048]
$\sigma(r_m - r_f)$	0.199 (0.021)	0.106 [0.041,0.162]	$sd(\Delta d)$	0.117 (0.020)	0.062 [0.025,0.090]
$E(p/d)$	3.377 (0.082)	2.951 [2.741,3.144]	$\rho_{\Delta c,\Delta d}$	0.59 (0.26)	0.67 [0.35,0.86]
$\sigma(p/d)$	0.450 (0.054)	0.195 [0.082,0.300]			
$E(R_s - R_b)$	0.094 (0.043)	0.083			
$E(R_v - R_g)$	0.073 (0.027)	0.037			

The table reports GMM estimates (asymptotic standard errors in parentheses) of the model parameters defined in Section 2.1. It also reports the median (95% confidence interval in square brackets), obtained through 10000 simulations, and the historical values (asymptotic standard errors in parentheses) of the mean and volatility of the risk free rate, price-dividend ratio, equity, size, and value premia, and unconditional moments of the consumption and dividend growth rates.

Table II: Summary Statistics in the Two Regimes, 1931-2009

	Panel A: Regime 1				*Panel B: Regime2*			
	Data		*Model*		*Data*		*Model*	
	$E(.)$	$sd(.)$	$E(.)$	$sd(.)$	$E(.)$	$sd(.)$	$E(.)$	$sd(.)$
Δc	0.016	0.017	0.013	0.020	0.007	0.044	0.030	0.033
	(0.002)	(0.002)	[−0.002,0.029]	[0.011,0.031]	(0.011)	(0.008)	[0.007,0.051]	[0.018,0.047]
Δd	0.036	0.067	−0.018	0.051	−0.063	0.204	0.044	0.069
	(0.011)	(0.008)	[−0.068,0.037]	[0.026,0.081]	(0.047)	(0.038)	[−0.026,0.103]	[0.039,0.102]
Δgdp	0.040	0.032			0.019	0.064		
	(0.005)	(0.004)			(0.017)	(0.007)		
$Inflation$	0.031	0.032			0.018	0.084		
	(0.005)	(0.004)			(0.024)	(0.016)		
$\log(P/D)$	3.474	0.432	2.937	0.177	3.047	0.349	3.086	0.208
	(0.090)	(0.056)	[2.751,3.131]	[0.090,0.284]	(0.092)	(0.060)	[2.848,3.295]	[0.107,0.331]
r_f	0.020	0.021	0.009	0.015	−0.033	0.088	0.028	0.020
	(0.003)	(0.002)	[−0.007,0.026]	[0.007,0.025]	(0.026)	(0.020)	[0.007,0.047]	[0.011,0.031]
r_m	0.053	0.185	0.036	0.095	0.111	0.233	0.091	0.150
	(0.023)	(0.018)	[−0.004,0.084]	[0.047,0.149]	(0.054)	(0.058)	[0.017,0.159]	[0.070,0.218]
$r_m - r_f$	0.033	0.182	0.027	0.093	0.144	0.235	0.063	0.153
	(0.022)	(0.017)	[0.001,0.061]	[0.044,0.150]	(0.059)	(0.068)	[0.005,0.118]	[0.068,0.224]
$r_s - r_b$	0.007	0.210			0.182	0.253		
	(0.033)	(0.022)			(0.067)	(0.024)		
$r_v - r_g$	0.024	0.201			0.116	0.218		
	(0.029)	(0.018)			(0.046)	(0.030)		

Panel A reports the sample mean and standard deviation (asymptotic standard errors in parentheses) of consumption, dividend, and GDP growth rates, the rate of inflation, log price-dividend ratio, risk free rate, market return, and equity, size, and value premia in the first regime. It also reports the median (95% confidence intervals in square brackets) of the mean and standard deviation of consumption and dividend growth rates, the log price-dividend ratio, risk free rate, market return, and equity premium in the first regime, obtained through 10000 simulations. Panel B reports the corresponding moments in the second regime.

Table III: Forecastability in the Two Regimes, 1931-2009

	Panel A: Regime 1		
	const.	*log* (P/D)	*Adjusted-*R^2
Δc	0.005 (0.019)	0.003 (0.005)	−0.010
Δd	0.054 (0.075)	−0.005 (0.021)	−0.016
r_m	0.332 (0.205)	−0.080 (0.058)	0.014
$r_m - r_f$	0.309 (0.201)	−0.079 (0.057)	0.015

	Panel B: Regime 2		
	const.	*log* (P/D)	*Adjusted-*R^2
Δc	−0.281 (0.081)	0.099 (0.027)	0.413
Δd	−1.136 (0.430)	0.367 (0.146)	0.237
r_m	−0.016 (0.580)	0.043 (0.197)	−0.059
$r_m - r_f$	−0.357 (0.572)	0.172 (0.195)	−0.013

Panel A reports regression coefficients (standard errors in parentheses) and adjusted-R^2 of in-sample linear regressions of the consumption and dividend growth rates and equity premium on the log price-dividend ratio as predictive variable in the first regime. Panel B reports the corresponding results in the second regime.

Table IV: Forecastability of the Equity, Size, and Value Premia, 1931-2009

Panel A: Equity Premium

const.	x	p	xp	$log\,(P/D)$	r_f	$log\,(P/D)\,r_f$	*Adjusted-*R^2
0.13 (0.05)	−0.80 (0.83)	-0.10^{*} (0.06)	1.79 (1.81)				0.074
0.36 (0.17)				-0.09^{*} (0.05)			0.028
0.33 (0.17)				−0.08 (0.05)	−0.60 (0.45)		0.038
0.40 (0.19)				-0.10^{*} (0.06)	−3.41 (3.61)	0.96 (1.23)	0.034

Panel B: Size Premium

const.	x	p	xp	$log\,(P/D)$	r_f	$log\,(P/D)\,r_f$	*Adjusted-*R^2
0.20 (0.05)	1.13 (0.91)	-0.21^{***} (0.06)	-5.45^{***} (1.98)				0.143
0.28 (0.20)				−0.07 (0.06)			0.006
0.25 (0.19)				−0.06 (0.06)	-0.86^{*} (0.52)		0.028
0.02 (0.21)				0.02 (0.06)	9.11^{**} (4.07)	-3.42^{**} (1.38)	0.089

Panel C: Value Premium

const.	x	p	xp	$log\,(P/D)$	r_f	$log\,(P/D)\,r_f$	*Adjusted-*R^2
0.15 (0.05)	1.43^{*} (0.86)	-0.14^{**} (0.06)	-2.69^{*} (1.87)				0.048
0.17 (0.18)				−0.04 (0.05)			−0.007
0.16 (0.18)				−0.03 (0.05)	−0.06 (0.48)		−0.020
0.07 (0.20)				−0.003 (0.06)	4.25 (3.85)	−1.48 (1.31)	−0.016

Panels A, B, and C report results from forecasting regressions for the equity, size, and value premia, respectively. The first row of each panel reports the regression coefficients along with the associated standard errors in parentheses, and the adjusted-R^2 from the forecasting regression of the realized premium on x, p, and xp. The second, third, and fourth rows report, respectively, the corresponding results when the set of predictor variables consists of the lagged aggregate log price-dividend ratio, the price-dividend ratio and log risk free rate, and the price-dividend ratio, risk free rate, and their product.

Table V: In- and Out-of-Sample Forecastability of Equity, Size, and Value Premia

Panel A: Equity Premium, 1976-2009

const.	x	p	xp	$log\,(P/D)$	r_f	$log\,(P/D)\,r_f$	*Adjusted-*R^2	R^2_{OOS}
0.09	2.01	−0.02	1.32				0.085	0.052
(0.06)	(2.36)	(0.07)	(3.77)					
0.36				−0.08			0.022	−0.046
(0.23)				(0.06)				
0.35				−0.08	0.06		−0.010	−0.025
(0.25)				(0.07)	(1.10)			
0.80				-0.21^{***}	-20.5^{**}	5.82^{**}	0.095	−0.029
(0.31)				(0.08)	(9.64)	(2.72)		

Panel B: Size Premium, 1976-2009

const.	x	p	xp	$log\,(P/D)$	r_f	$log\,(P/D)\,r_f$	*Adjusted-*R^2	R^2_{OOS}
0.10	−1.57	−0.09	-6.51^{*}				0.255	0.226
(0.06)	(2.37)	(0.07)	(3.78)					
−0.03				0.01			−0.030	−0.149
(0.29)				(0.08)				
0.19				−0.02	-3.39^{***}		0.155	−0.068
(0.27)				(0.07)	(1.20)			
0.08				0.006	1.50	−1.39	0.133	0.053
(0.37)				(0.099)	(11.2)	(3.17)		

Panel C: Value Premium, 1976-2009

const.	x	p	xp	$log\,(P/D)$	r_f	$log\,(P/D)\,r_f$	*Adjusted-*R^2	R^2_{OOS}
0.15	3.45	−0.09	-8.49^{**}				0.059	−0.003
(0.06)	(2.59)	(0.08)	(4.14)					
0.22				−0.04			−0.022	−0.114
(0.28)				(0.08)				
0.29				−0.05	−1.05		−0.033	−0.119
(0.29)				(0.08)	(1.29)			
−0.09				0.05	16.4	−4.96	0.008	−0.103
(0.38)				(0.10)	(11.7)	(3.30)		

Panels A, B, and C report in-sample and out-of-sample forecasting results for the equity, size, and value premia, respectively. The first row of each panel reports the in-sample regression coefficients along with the standard errors in parentheses, and the adjusted-R^2 from the forecasting regression of the realized premium on x, p, and xp. It also reports the out-of-sample R^2 from rolling predictive regressions on x, p, and xp. The second, third, and fourth rows report, respectively, the corresponding results when the set of predictor variables consists of the lagged aggregate log price-dividend ratio, the price-dividend ratio and log risk free rate, and the price-dividend ratio, risk free rate, and their product.

Table VI: Forecastability of the Consumption Growth Rate

Panel A: Consumption Growth, 1931-2009						
$const.$	x	p	$log(P/D)$	r_f	$Adjusted - R^2$	
0.004 (0.006)	-0.233^{***} (0.088)	0.012 (0.007)			0.080	
-0.039 (0.021)			0.016^{***} (0.006)		0.068	
-0.043 (0.021)			0.017^{***} (0.006)	-0.086 (0.056)	0.084	
Panel B: Consumption Growth, 1947-2009						
$const.$	x	p	$log(P/D)$	r_f	$Adjusted - R^2$	
0.006 (0.006)	0.302^{***} (0.089)	0.008 (0.006)			0.214	
-0.025 (0.016)			0.011^{**} (0.005)		0.067	
-0.014 (0.015)			0.007^{*} (0.004)	0.201^{***} (0.052)	0.239	
Panel C: Consumption Growth, 1976-2009						
$const.$	x	p	$log(P/D)$	r_f	$Adjusted - R^2$	R^2_{OOS}
0.008 (0.006)	0.346^{***} (0.123)	0.007 (0.007)			0.153	0.025
-0.009 (0.020)			0.006 (0.005)		0.009	-0.959
-0.025 (0.018)			0.009^{*} (0.005)	0.264^{***} (0.081)	0.236	-1.232

Panels A and B report in-sample forecasting results for consumption growth over 1931–2009 and 1947-2009, respectively. Panel C reports in-sample forecasting and out-of-sample predictive results over 1976-2009.

Table VII: Forecastability of the Dividend Growth Rate

Panel A: Dividend Growth, 1931-2009

$const.$	x	p	$log(P/D)$	r_f	$Adjusted-R^2$
-0.071 (0.027)	-0.347 (0.401)	0.111^{***} (0.032)			0.117
-0.256 (0.097)			0.080^{***} (0.029)		0.080
-0.251 (0.099)			0.078^{***} (0.029)	0.119 (0.263)	0.070

Panel B: Dividend Growth, 1947-2009

$const.$	x	p	$log(P/D)$	r_f	$Adjusted-R^2$
0.013 (0.030)	1.078^{**} (0.448)	0.015 (0.032)			0.087
-0.002 (0.078)			0.007 (0.022)		-0.015
0.037 (0.076)			-0.007 (0.022)	0.738^{***} (0.269)	0.084

Panel C: Dividend Growth, 1976-2009

$const.$	x	p	$log(P/D)$	r_f	$Adjusted-R^2$	R^2_{OOS}
-0.000 (0.024)	0.536 (0.734)	0.041 (0.032)			0.029	0.044
-0.022 (0.114)			0.013 (0.031)		-0.026	-0.602
-0.089 (0.113)			0.025 (0.030)	1.068 (0.502)	0.076	-0.603

Panels A and B report in-sample forecasting results for dividend growth over 1931–2009 and 1947-2009, respectively. Panel C reports in-sample forecasting and out-of-sample predictive results over 1976-2009.

Table VIII: Forecastability of Variance of Market Return and Growth Rates, 1931-2009

const.	x	p	xp	$log\,(P/D)$	r_f	$log\,(P/D)\,r_f$	*Adjusted-*R^2
Panel A: Variance of Market Return							
0.003 (0.001)	0.016 (0.012)	-0.002^{**} (0.001)	-0.031 (0.026)				0.021
0.005 (0.002)				-0.001 (0.001)			0.010
0.005 (0.002)				-0.001 (0.001)	0.003 (0.006)		0.001
0.005 (0.003)				-0.001 (0.001)	-0.022 (0.052)	0.009 (0.018)	-0.009
Panel B: Variance of Consumption Growth, $(\Delta c_{t+1} - E\left[\Delta c_{t+1} \vert \mathcal{F}\,(t)\right])^2$							
const.	x	p	xp	$log\,(P/D)$	r_f	$log\,(P/D)\,r_f$	*Adjusted-*R^2
0.0014 (0.0003)		-0.0011^{***} (0.0003)					0.037
0.003 (0.001)				-0.0006 (0.0004)			0.027
0.002 (0.001)				-0.0005^{*} (0.0003)	0.0008 (0.0028)		0.006
Panel C: Variance of Dividend Growth, $(\Delta d_{t+1} - E\left[\Delta d_{t+1} \vert \mathcal{F}\,(t)\right])^2$							
const.	x	p	xp	$log\,(P/D)$	r_f	$log\,(P/D)\,r_f$	*Adjusted-*R^2
0.036 (0.007)		-0.032^{***} (0.008)					0.155
0.061 (0.025)				-0.014^{**} (0.007)			0.034
0.062 (0.025)				-0.015^{**} (0.007)	0.016 (0.068)		0.024

Panels A, B, and C report results of forecasting regressions for the varianc of the market return and consumption and dividend growth rates, respectively.

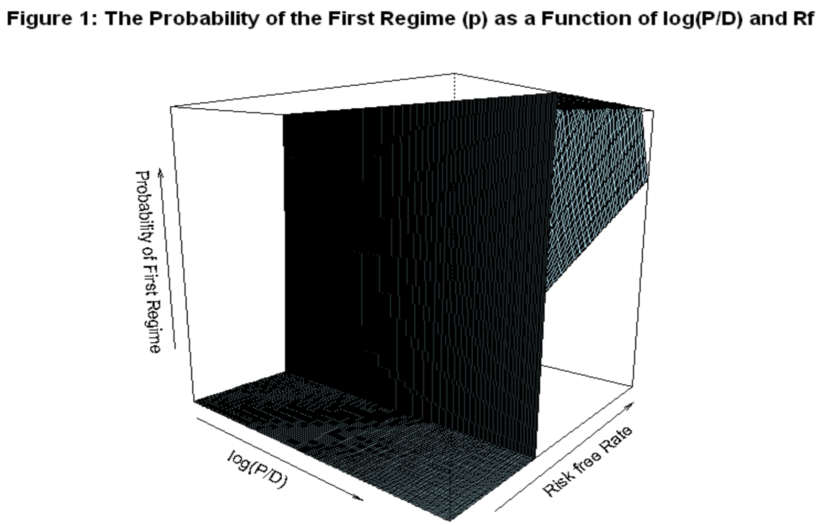

Figure 1: The figure plots the probability of being in the first regime against the price-dividend ratio and risk free rate.

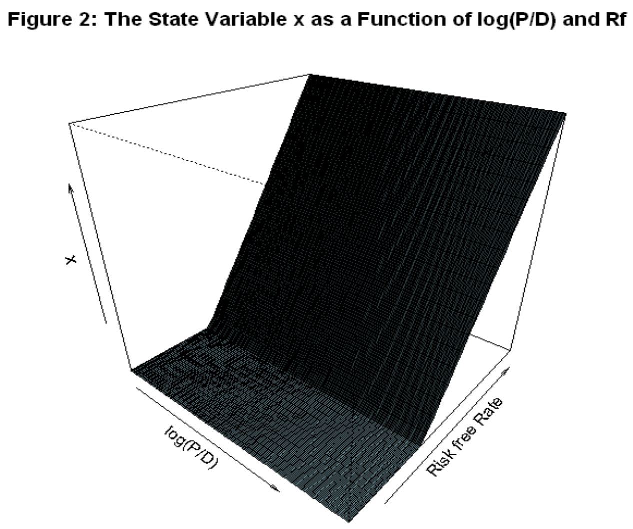

Figure 2: The figure plots the state variable x against the price-dividend ratio and risk free rate.

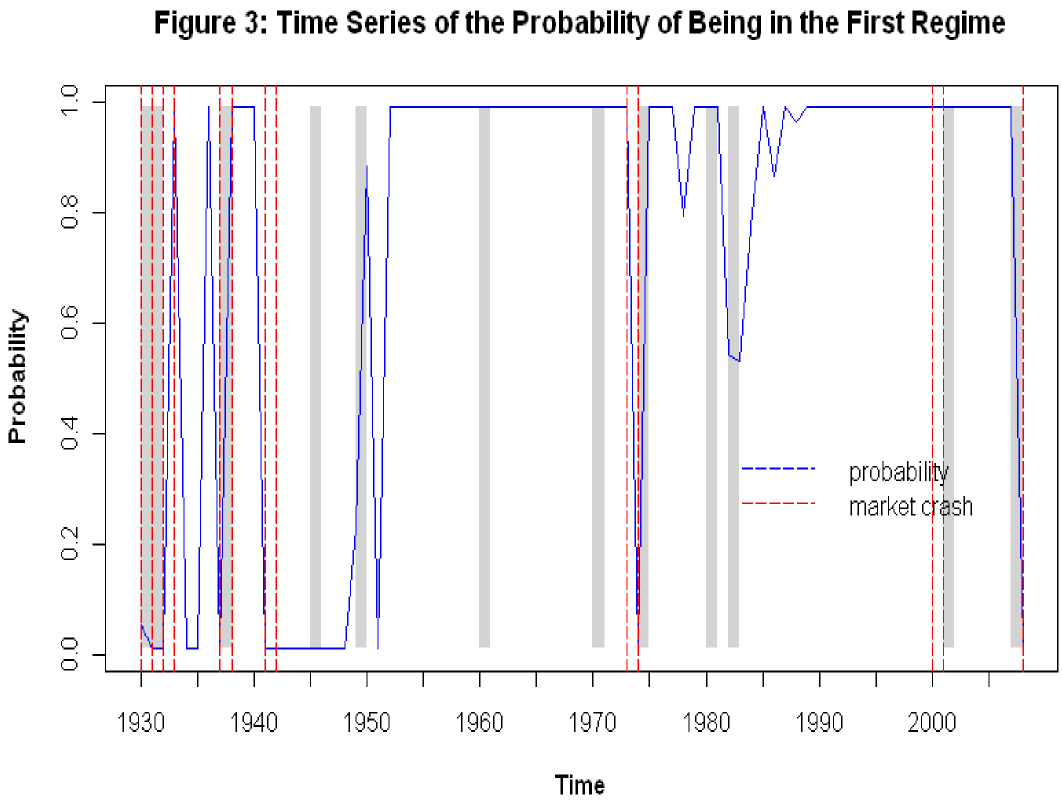

Figure 3: The Figure presents the time series of the probability of being in the first regime over $1930 - 2008$ along with the NBER recessions (shaded columns) and the major stock market downturns (vertical dashed lines) identified in Barro and Ursua (2009).

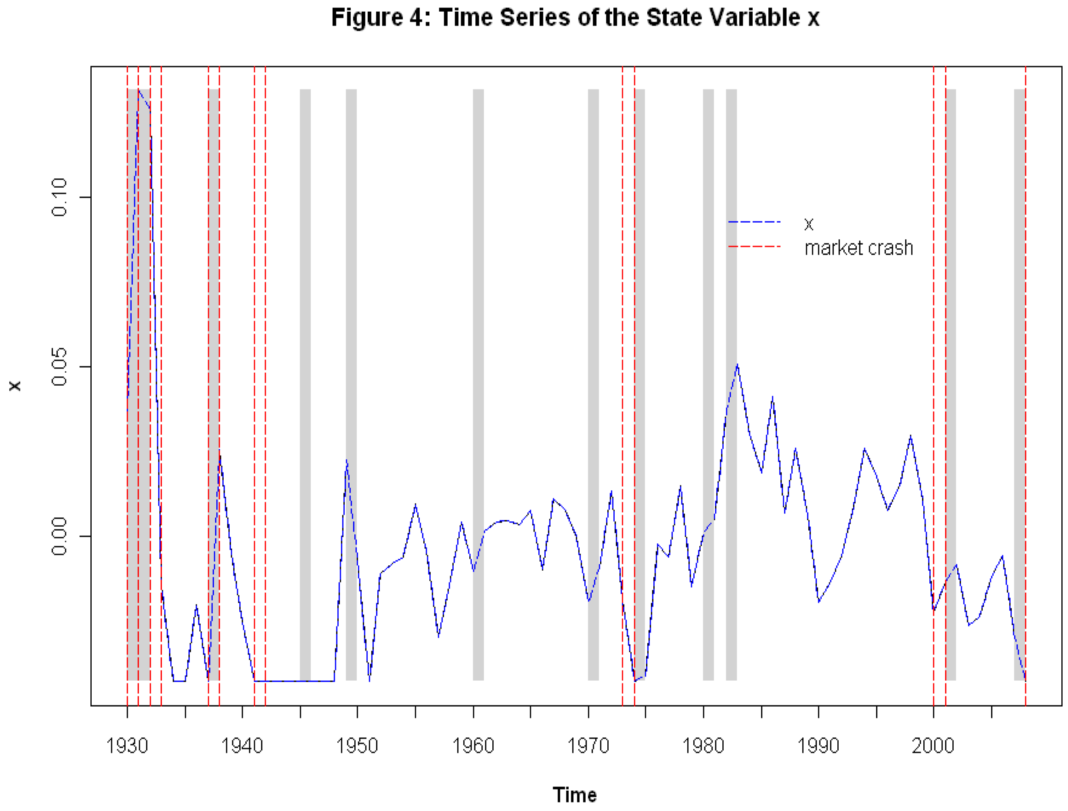

Figure 4: The Figure presents the time series of the state variable x over $1930 - 2008$ along with the NBER recessions (shaded columns) and the major stock market downturns (vertical dashed lines) identified in Barro and Ursua (2009).

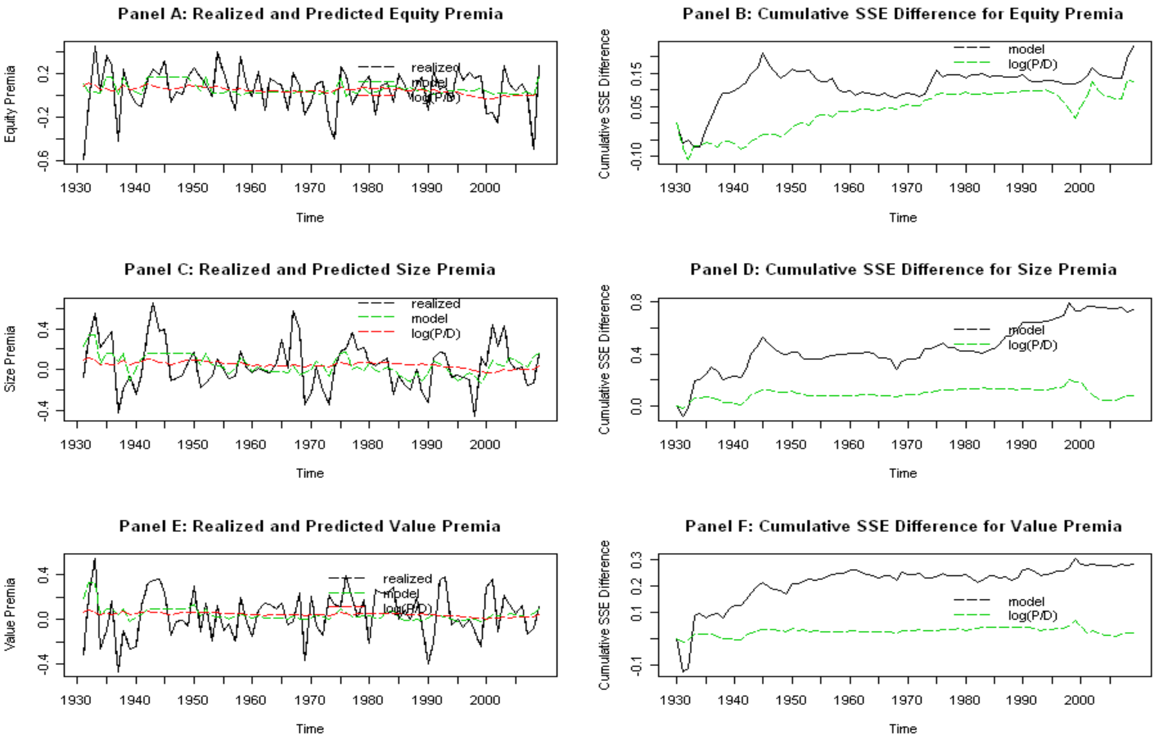

Figure 5: Panels A, C, and E display the realized equity, size, and value premia (black solid line), respectively, along with their forecasted values from the regressions implied by the regime shift model (green dotted line) and linear regressions using the log market-wide price-dividend ratio as a predictor variable (red dashed line). Panels B, D, and F display the cumulative squared demeaned equity, size, and value premia, respectively, minus the cumulative squared regression residual from the alternative forecasting regression specifications: the forecasting regression implied by the model (black solid line) and a linear forecasting regression with the log price-dividend ratio as a predictor variable (red dashed line). An increase in a line indicates better performance of the named model relative to the historical mean of the premia while a decrease in a line indicates better performance of the historical mean.

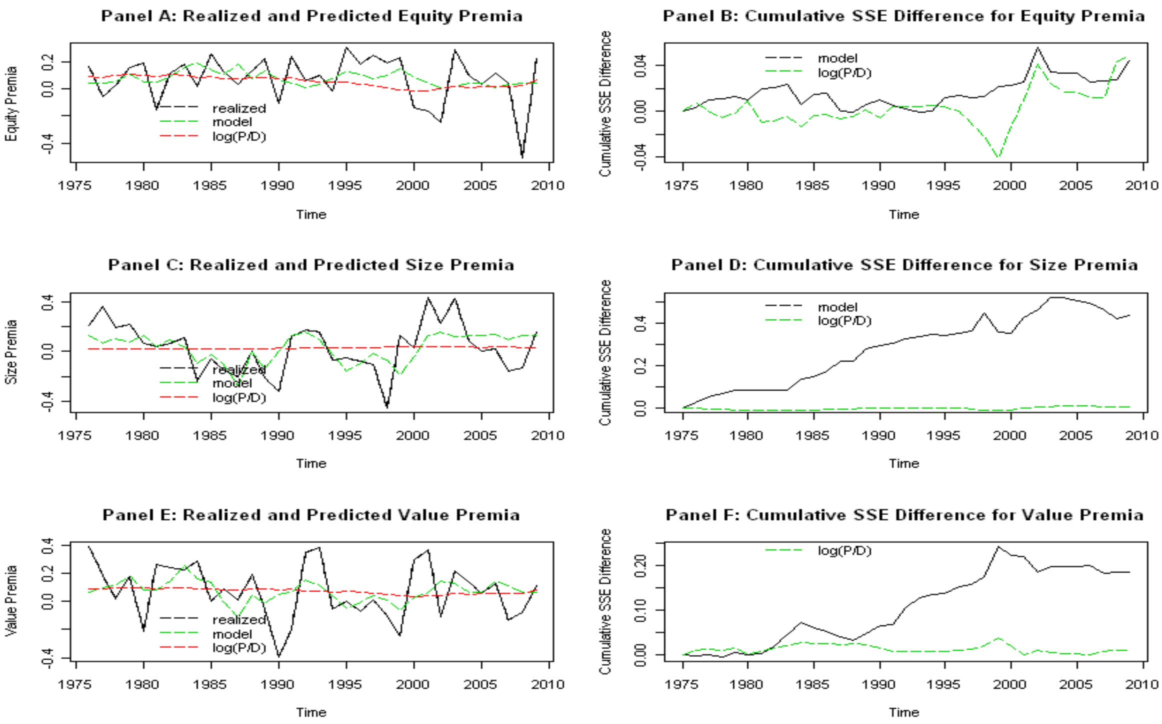

Figure 6: Panels A, C, and E display the realized equity, size, and value premia (black solid line), respectively, along with their forecasted values from the regressions implied by the regime shift model (green dotted line) and linear regressions using the log market-wide price-dividend ratio as a predictor variable (red dashed line). Panels B, D, and F display the cumulative squared demeaned equity, size, and value premia, respectively, minus the cumulative squared regression residual from the alternative forecasting regression specifications: the forecasting regression implied by the model (black solid line) and a linear forecasting regression with the log price-dividend ratio as a predictor variable (red dashed line). An increase in a line indicates better performance of the named model relative to the historical mean of the premia while a decrease in a line indicates better performance of the historical mean.

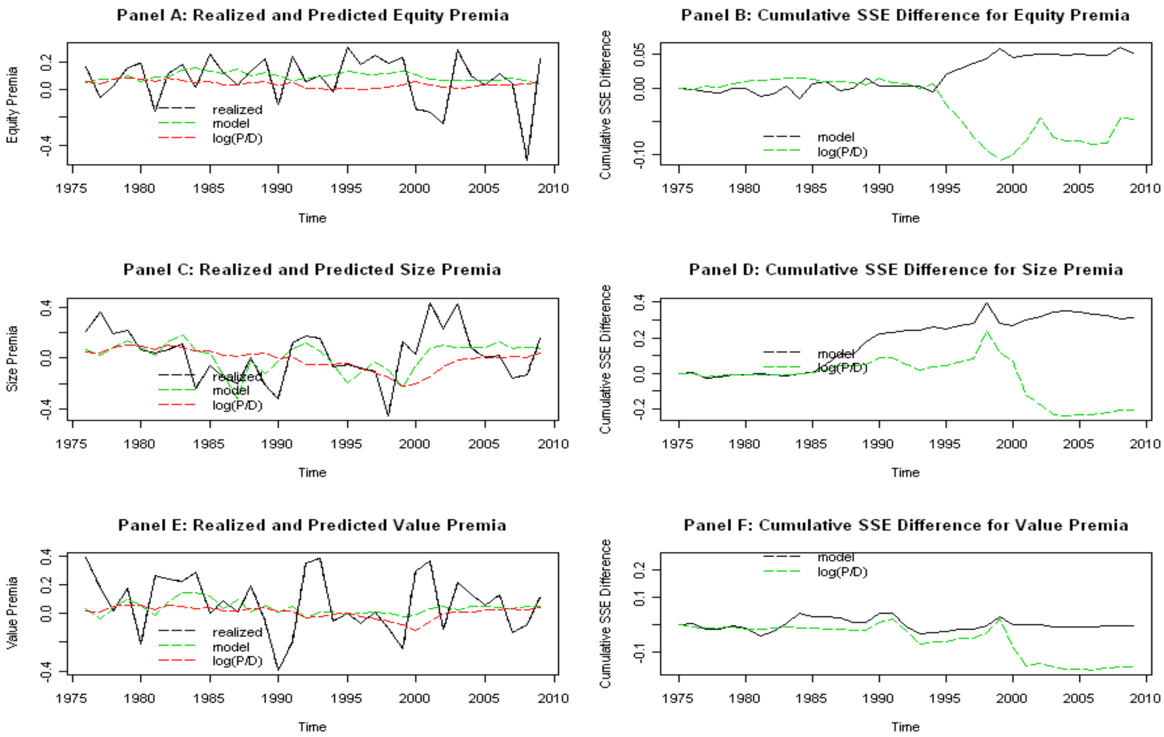

Figure 7: Panels A, C, and E display the realized equity, size, and value premia (black solid line), respectively, along with their predicted values from the rolling out-of-sample predictive regressions implied by the regime shift model (green dotted line) and linear regressions using the log market-wide price-dividend ratio as a predictor variable (red dashed line). Panels B, D, and F display the cumulative squared demeaned equity, size, and value premia, respectively, minus the cumulative squared regression residual from the alternative forecasting regression specifications: the forecasting regression implied by the model (black solid line) and a linear forecasting regression with the log price-dividend ratio as a predictor variable (red dashed line). An increase in a line indicates better performance of the named model relative to the historical mean of the premia while a decrease in a line indicates better performance of the historical mean.

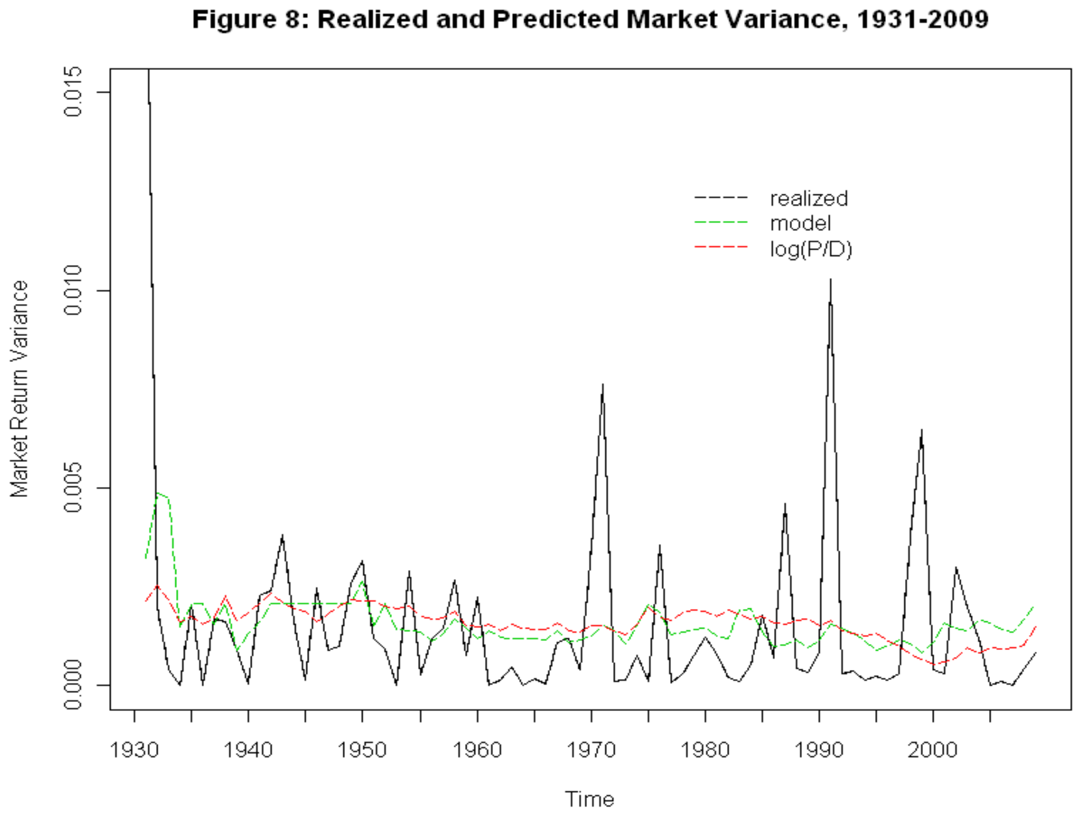

Figure 8: The figure plots the realized market variance along with its model-predicted value. The black solid line plots the annual realized variance that is calculated as the sum of squares of the monthly log returns. The green dotted line plots the model-predicted variance and is obtained as the fitted value from a regression of the realized variance on the two state variables and their product. The red dashed line shows the predictive power of the log price-dividend ratio by plotting the fitted value from a regression of the realized variance on the lagged log price-dividend ratio.