Equilibrium Predictability And Other Return Characteristics

Abstract

This paper presents a structural model of aggregate return characteristics based on a one-channel Bansal and Yaron (2004) economy under recursive preferences. The predictability result rests on an endogenously determined price-dividend ratio that is not exponentially affine which implies time variation and predictability in equity premia. This result is new within the context of a one-channel Bansal and Yaron (2004) economy. Furthermore, the predictability coefficient is stochastic which provides theoretical foundations for recent works in predictability like Dangl and Halling (2011). In longer horizon, the predictability relationship is highly volatile making it difficult to make inference about long-horizon predictability.

Introduction

This paper investigates structural return predictability within a one-channel long-run risk framework of Bansal and Yaron (2004). Starting from simple joint dynamics of aggregate consumption and dividend growth, where the expected growth rates of both share a common stochastic trend, I find that equity premium is time-varying. This finding is significant because Bansal and Yaron (2004) does not have time-varying equity premium in their one-channel economy. A natural consequence of this is time-variation in the coefficient of return predictability. In my model, dividend yield and the coefficient of return predictability are inversely related - a decrease in dividend yield corresponds to a rapid increase in the return predicting coefficient which increases equity premium. Furthermore, the setting is tractable enough to solve for the long-horizon regression predictability coefficient in semi-closed form. The time-series of long-horizon predictability coefficients show significant variation in US data. This time-variation leads to high variance in return predictability, especially over longer horizon, making long-horizon predictability of returns extremely unreliable.

The theory developed here makes two potentially interesting points. First, it finds time-series implications in equity premium in the Bansal and Yaron (2004) one channel economy. The reason that Bansal and Yaron (2004) does not find time-variation in expected returns in their one-channel economy is because they *posit* price-dividend (PD) ratio to be exponentially affine. Thus, the volatility of PD ratio is a constant, and, since the one-channel economy has constant volatility in dividend and consumption growth, equity premium is constant and there is no dynamic return predictability. In their work, the dual channel economy with stochastic volatility of the growth rate of consumption and dividend creates time-variation in the volatility of the pricing kernel which gives rise to time-varying risk-premia and predictability. In this paper, I start with Duffie-Epstein preferences with elasticity of intertemporal substitution equal to unity, which allows me to solve for the Hamilton-Jacobi-Bellman equation of the representative agent in closed form. Subsequently, this gives me a closed form expression of the pricing kernel which allows me to get an analytical expression for the PD ratio that is no longer exponentially affine. The non-linearity in the log PD ratio creates time-varying volatility in returns, which gives rise to time-varying risk-premia and predictability. The non-linearity in the PD ratio implies that the quantity of risk is time-varying. In response to a good expected dividend growth shock, the agent buys more of the asset that pays aggregate dividends which increases the quantity of risk that the agent bears. The opposite happens in response to a bad shock. Furthermore, in the Appendix I relaxed the assumption of unit EIS and found that the above dynamics of PD ratio is consistent with EIS greater than unity. Additionally, equity-premium is positive and pro-cyclical as long as risk-aversion is greater than the inverse of EIS, i.e. as long as the agent has preference for early resolution of uncertainty. The economic effect of this non-linearity in the PD ratio manifests in return predictability, which is the second major point in the paper.

Given the closed form solution of PD ratio, I can directly solve for the long-horizon return predictability coefficient in a semi-closed form. This coefficient is time-varying and reflects at any time t, the agent's expectation of price and dividend growth over a particular horizon. The time-variation in predictability coefficients is yet unexplored in the equilibrium literature although it has gained significant attention in the empirical works of Lettau and Van Nieuwerburgh (2008) and Dangl and Holling (2011). These recent works on predictability show substantial uncertainty in estimating the predictability coefficient and Dangl and Holling (2011) address that by modelling the coefficient in a state-space framework. The time-varying coefficient significantly helps out-of-sample forecasts, and investors armed with such models outperform investors with constant coefficient models. My model provides theoretical foundation for time-varying predictability coefficients. An immediate conclusion from my model is that OLS regressions for predictability can be deeply misspecified and the parameter uncertainty that these papers encounter is precisely due to the time-variation in the predictability coefficient derived here.

This paper rests on the non-linearity in the log PD ratio. Is the level of non-linearity in the log PD ratio significant? Let's assume that it is not very significant. Therefore, the volatility of PD ratio should be a constant[1], and since the volatility of the pricing kernel is constant, expected return should also be constant across different realizations of the latent growth rate. However, the qualitative and quantitative properties of the model are quite different. The top panel of Figure 3 shows that changes in equity premium across different states of the latent growth rate is quite significant, thus confirming the economic siginificance of the non-linearity.

Long-horizon predictability has received a lot of attention in the empirical literature since the early studies of Shiller (1981), Rozeff (1984), Campbell and Shiller (1988), Fama and French (1988), among others. Fama and French (1988) were the first to show long-horizon predictability reporting coefficients and R^2-s that increase with horizon. Since then, others have cast doubt on long-horizon predictability. Stambaugh (1999) finds severe biases in small sample estimators. Although he doesn't look at long-horizon regressions, his criticism applies to long-horizon regressions as well. In finite samples, he finds that p-values for the coefficients are much higher than OLS p-values casting doubt on the significance of the coefficient. For long-horizon predictability, Valkanov (2003) shows that the coefficients have limiting distributions that are functionals of Brownian shocks and the OLS

[1] volatility of dividend growth is constant in the one channel economy

estimators of them are highly inconsistent. Goetzmann and Jorion (1993) find evidence in their simulation based study that if returns follow a random walk but dividends are autocorrelated, then the R^2's from long-horizon regressions could still be significant even though there is no underlying predictability. For reasons similar to Valkanov (2003), they conclude that the right-hand side variables are strongly correlated to lagged left-hand side variables which gives rise to these fictitous R^2's when there is no underlying predictability. Recently, Boudoukh, Richardson and Whitelaw (2008) (BRW) show that there is no extra information in long-horizon regressions than is already factored in short-horizon ones. In fact, they show that their return predicting coefficients and R^2's, when represented as multiples of one-year coefficient or R^2's, scale perfectly with time. This paper follows in the same spirit as Goetzmann and Jorion (1993) modeling underlying autocorrelated shocks in the dividends that show up in both dividend yields and returns, thereby creating the environment where underlying shocks can build through time to produce high predictability coefficients.

Other equilibrium asset pricing models like Campbell and Cochrane (1999) and Bansal and Yaron (2004) do not investigate time-variation in predictability coefficients. These models simulate long-horizon returns and run OLS regressions to show evidence of long-horizon predictability characterized by increasing coefficients and R^2s. I replicate these regressions in my model and get the same results. However, in my model, I am able to solve for the predictability coefficient and without running any regressions I can evaluate quantitative and qualitative properties of the coefficients. The coefficients show significant time-variation, especially in longer horizon. However, this time-variation comes with a price. To gauge what kind of unconditional inference can be drawn, I summarize the variance of long-horizon predictability relationship relative to the variance of long-horizon returns in a pseudo-R^2 quantity. This pseudo-R^2 increases over the horizon purely due to the fact that the variance of the predictability component increases faster than the variance of overall returns. Thus, over longer horizon this pseudo-R^2 increases artificially, whereas, at least qualitatively it is hard to justify predictive power when this increase in pseudo-R^2 is due to the increase in variance of the predictability component. In essence, my equilibrium model tells the following story - the parameters that comply with macroeconomic dynamics and can match key asset pricing quantities imply significant time-variation in predictability coefficients. This time-variation implies large unconditional variances of the predictable component of long-horizon returns which creates doubt about return predictability in the long horizon.

This paper also fits into the growing body of literature that endogenizes prices to address the question of predictability like Menzly, Santos and Veronesi (2004) and Ang and Liu (2007). Menzly, Santos and Veronesi (2004) deal with predictability in the cross-section under habit persistance. Ang and Liu (2007) show the implication of endogenizing any two of expected returns, dividend yields and volatilities, given any one of them and dividend growth dynamics. Unlike Menzly, Santos and Veronesi (2004), I focus on predictability from the perspective of long-run risk induced by a time-varying fluctation of expected growth rates and DE preferences, and unlike Ang and Liu (2007) I endogenize all three of expected returns, volatility and dividend yields given dividend dynamics and DE preferences.

The paper is subdivided into the following parts: section 1 discusses the details of the model and establishes the predictability results. Section 2 discusses the estimation methodology. Section 3 covers the empirical findings on asset pricing quantities and equilibrium predictability.

1 The Model

1.1 Preferences and Dynamics

Power utility puts a heavy restriction on risk-aversion and elasticity of intertemporal substitution (EIS)- they are reciprocals of each other. EIS measures willingness to exchange non-stochastic consumption today for tomorrow given a particular interest rate today. As such, the restriction that power utility imposes is too strict on two very different concepts - risk aversion is about preference over a random variable and EIS is substitution across deterministic consumption paths. In equilibrium asset pricing, the power utility restriction amounts to jointly establishing both the risk-free rate and equity premium through the same parameter - risk aversion. Empirically, the power utility restriction is a dismal failure giving rise to the equity premium puzzle and the corresponding risk-free rate puzzle. To break the strict relationship between the two, recursive utility functions are introduced a la Epstein-Zin-Weil that considers the concepts separately.

The utility function that is considered here is due to Duffie and Epstein (1992) which is a continuous time counterpart of Kreps-Porteus and Epstein-Zin-Weil preferences. The normalized utility function considered here is

$$f(C,J) = \frac{\beta(1-\gamma)J}{1-\frac{1}{\psi}}\left[C^{1-\frac{1}{\psi}}\left((1-\gamma)J\right)^{\frac{\frac{1}{\psi}-1}{1-\gamma}} - 1\right]$$

where C is the current period consumption, J is the value function, ψ is the EIS, β is the discount rate and γ is the risk-aversion. Assume furthermore that the representative investor is endowed with a log-recursive utility, which is a special case of the above preference with $\psi = 1$. The above utility function simplifies substantially in the $\psi = 1$ special case to

$$f(C, J) = \beta(1-\gamma)J\left[\log C - \frac{\log(1-\gamma)J}{1-\gamma}\right]$$

The appendix also solves the model for the general case using log-linearization.

Assume that consumption and dividend growth jointly follow a geometric path with mean reverting growth rate X_t,

$$\begin{aligned}
\frac{dD}{D} &= (\mu_D + X_t)dt + \sigma_D dW_D && (1)\\
\frac{dC}{C} &= (\mu_C + \lambda X_t)dt + \sigma_C dW_C && (2)\\
dX_t &= -\kappa X_t dt + \sigma_x dW_X && (3)
\end{aligned}$$

where the Brownian motion shocks are all uncorrelated. This formulation has its origin in Abel (1999) and this is very similar to the one-channel model of Bansal and Yaron (2004) except for one caveat - the parameter λ loads on the latent shock in consumption growth rate instead of dividends. When a growth rate shock jointly hits expected dividend and consumption growth, $\lambda < 1$ has the effect of tempering down the corresponding expected consumption growth rate relative to dividend growth rate. This fact is also borne out in the data. Figure 1 shows that real dividend growth rate has a lot of time-series variation whereas the corresponding real consumption growth is quite smooth, and $\lambda < 1$ helps us achieve that.

More interestingly, notice that the volatility of dividend and consumption growth are non-stochastic and what I will show below is that unlike Bansal and Yaron (2004) one-channel economy it still produces time-varying risk-premium. Furthermore notice that the correlation between all the Brownian motion shocks is set to zero, so that I can devote the full attention to market price of risk and risk-premia stemming from the long-run risk due to growth rate X_t.

The utility process J satisfies the Bellman equation with respect to equilibrium consumption

$$\mathcal{DJ}(C, X, t) + f(C, J) = 0 \tag{4}$$

where $\mathcal{DJ}$ is the differential operator applied to J with respect to $\{C, X, t\}$ with the boundary condition $J(C, X_T, T) = 0$. I am interested in the equilibrium as $T \to \infty$. Thus, I drop the explicit time dependence assuming that the agent is infinilitely long-lived and has reached equilibrium over time.

Proposition 1.1.1 *The solution to the Bellman equation in (4) corresponding to growth rate dynamics in (1)-(3) and preferences given by Duffie-Epstein utility with EIS=1 is*

$$J(C_t, X_t) = \frac{C_t^{1-\gamma}}{1-\gamma} \exp\left(u_1 X_t + u_2\right) \tag{5}$$

where

$$\begin{aligned} u_1 &= \frac{\lambda(1-\gamma)}{\kappa+\beta} \\ u_2 &= \frac{1-\gamma}{\beta}\left[\mu_C - \frac{1}{2}\gamma\sigma_C^2 + \frac{\lambda^2(1-\gamma)\sigma_x^2}{2(\kappa+\beta)^2}\right] \end{aligned}$$

1.2 Asset Pricing

Duffie and Epstein (1992) show that the pricing kernel for stochastic differential utility, Λ_t, is given by $\Lambda_t = \exp(\int_0^t f_J ds) f_c$. It has a particularly nice and elegant expression in closed form for $\psi = 1$, and the appendix shows a version corresponding to the log-linearized solution of the value function for $\psi \neq 1$.

Proposition 1.2.1 *The pricing kernel for EIS=1 is given by*

$$\frac{d\Lambda}{\Lambda} = -r_t^f dt - \gamma \sigma_C dW_C - \frac{(\gamma - 1)\lambda}{\kappa + \beta} \sigma_x dW_X \tag{6}$$

where

$$r_t^f = \mu_C + \lambda X_t + \beta - \gamma \sigma_C^2 \tag{7}$$

The risk-free rate in (7) has many desirable properties which we do not observe in risk-free rate derived from standard power utility setting. In this case, risk-free rate is actually decreasing *uniformly* as risk-aversion, γ, increases, whereas in power utility I would need γ really high for the precautionary savings term to kick-in and generate the same effect. At that high level of risk-aversion, power utility implies that a one-percent increase in consumption growth would increase the risk-free rate by γ-percent - a claim not supported by the data. In the log-recursive case, a one-percent increase in consumption growth signifies a one-percent increase in the risk-free rate due to $\psi = 1$. The proposition that risk-free rate decreases in risk-aversion uniformly in the log-recrusive case is not surprising. Recall, that in the log-recursive case $\gamma > 1$ is sufficient to generate preference for early resolution of uncertainty. For high values of risk-aversion, i.e. strong preference for early resolution of uncertainty, the agent is willing to settle for a lower certainty equivalence in the future which reduces the risk-free rate. For $\psi > 1$ and higher preference for early resolution of uncertainty, the appendix shows that the risk-free rate falls sharply and is less responsive to changes in consumption growth rate than in the unit EIS case.

Since there are two sources of consumption risk in this economy there are two market prices of risk in (6). The first one is the traditional transient consumption risk term from power utility coming from volatility of consumption growth, and the second is due to the stochastic growth rate of consumption and recursive preferences and is popularly termed long-run risk. Notice that if $\lambda = 0$ and there was no stochastic growth rate of consumption then the long-run risk term will be zero. Moreover, notice that the long-run risk coefficient $\frac{(\gamma-1)\lambda}{\kappa+\beta}\sigma_x = \frac{J_X}{J}\sigma_x$ measures change in the value function of the agent with respect to the growth rate X_t. In recursive preferences, the value function is embedded within the utility function. Thus volatility in marginal utility necessarily measures volatility in the life-term utility of the agent - hence the name long-run risk.

Long-run risk is increasing in γ, but the effect is magnified due to κ and β in the denominator. Recall, that the stationary distribution of $X_t \sim N\left(0, \frac{\sigma_x}{\sqrt{2\kappa}}\right)$. Thus, as κ decreases and the growth rate becomes more persistant, the volatility of growth rate increases and an agent exposed to long-run risk from the volatile growth rate shocks seeks higher compensation for bearing this risk. Notice, that the magnitude of the size of long-run risk can be much higher vis-a-vis the risk from the transient consumption volatility as is shown in Table II.

The long-run market price of risk for $\psi \neq 1$ is directly proportional to $\gamma - \frac{1}{\psi}$, which distinguishes DE preferences from standard time-separable preferences where $\gamma = \frac{1}{\psi}$. Clearly, for time-separable preferences long-run risk vanishes. The quantity $\gamma - \frac{1}{\psi}$ also determines preference for early resolution of uncertainty. Thus, stronger the preference for early resolution of uncertainty of the growth rates the higher the market price of risk.

Given the pricing kernel of the stochastic differential utility, I establish the equilibrium price-dividend ratio and return dynamics.

Proposition 1.2.2 *Equilibrium price-dividend ratio is given by*

$$\frac{P_t}{D_t} = G(X_t) \tag{8}$$

where $G(X_t) = \int_t^\infty \exp(P_1(\tau)X_t + P_2(\tau))ds$, *where* $\tau = s - t$, $P_1(\tau)$ *and* $P_2(\tau)$ *are solutions of a system of ODEs given in the appendix. The dynamics for cumulative excess return is given by*

$$dR = \frac{dP + Ddt}{P} - r_t^f dt = \mu_t^R dt + \sigma_D dW_D + \frac{G_X}{G}\sigma_x dW_x$$

where equilibrium expected excess return is

$$\mu_t^R = \frac{\lambda(\gamma - 1)}{\kappa + \beta}\frac{G_X}{G}\sigma_x^2 \tag{9}$$

and the volatility of cumulative return given by

$$\sigma_t^R = \sqrt{\sigma_D^2 + \left(\frac{G_X}{G}\sigma_x\right)^2} \tag{10}$$

This is the central result in the paper. Unlike Bansal and Yaron (2004), the PD ratio in this one-channel economy is no longer exponentially affine but is non-linear in the expected growth rate. The non-linearity of the growth rate in the log PD ratio is responsible for generating time-varying equity premia and a dynamic predictability relationship. Notice, that in the Bansal and Yaron (2004) economy, the assumed PD ratio is exponentially affine in the growth rate (and also in conditional variance in the two-channel case) which makes conditional volatility of PD ratio a constant, i.e. if $G = \frac{P}{D} = \exp(a + bX_t)$, where a and b are constants, then Vol $\left(\frac{dG}{G}\right) = b\sigma_x$. Thus, if market price of risk from X_t is also a constant, then risk premium will be a constant thus eliminating any time-series phenomenon in expected returns. In the Bansal and Yaron (2004) economy there is time variation in equity premia only in the two-channel case due to time-varying volatility of consumption growth which creates time-varying market prices of risk. In this case, the market price of risk is constant, but the non-linearities in the log PD ratio creates time-varying volatilities in prices which creates time-varying equity premia. In other words, the market price of risk is constant but the quantity of risk is time-varying which gives rise to time-varying premia.

The cumulative return volatility (10) has two components - the first one is the transient risk of the volatility of dividend growth and the other is due to long-run risk. To reinforce the point on the non-linearity of the PD ratio, notice that the long-run risk component of volatility is time-varying precisely because $G(X_t)$ is not exponentially affine in the growth rate X_t, which ensures that $\frac{G_X}{G}$ would be time-varying making return volatility stochastic.

The expected excess return (9) seeks compensation for only long-run risk since the correlation between all the Brownian motion terms are shut off. It is straight-forward to incorporate those kind of risks from correlation, but for brevity I focus only on the long-run risk component arising from non-linearity in the PD ratio. Notice that

$$G_X = \frac{1-\lambda}{\kappa} \int_t^{\infty} \exp(P_1(\tau)X_t + P_2(\tau))(1 - e^{-\kappa\tau})d\tau \tag{11}$$

If $\lambda < 1$, as has been assumed in the model to make expected consumption growth "slower" than expected dividend growth, then $G_X > 0$ which guarantees that expected return is always positive. In the $\psi \neq 1$ case, $G_X > 0$ if $\psi > 1$ along with $\lambda < 1$. Moreover, expected return is positive and procyclical so long as the agent has preference for early resolution of uncertainty, i.e. $\gamma - \frac{1}{\psi} > 0$. Thus, the above results that are obtained in closed form for unit EIS carries over for $\psi \neq 1$, as long as γ and ψ are both greater than unity.

1.3 Time-Series and Predictability

The non-linearity in the log price-dividend ratio presents valuable time-series dynamics of aggregate returns. Before discussing predictability, it is essential that I discuss the time-series nature of expected return. Since $G_X > 0$, expected return is always positive. Moreover, since

$$\left(\frac{G_X}{G}\right)_X = \frac{1}{G^2}\left[\int_t^\infty \exp(\cdot)d\tau \int_t^\infty \exp(\cdot)P_1^2(\tau)d\tau - \left(\int_t^\infty \exp(\cdot)P_1(\tau)d\tau\right)^2\right] > 0$$

expected return is increasing in X_t.[2] Therefore, both PD ratio, and expected return rises with positive growth rate shock. In response to a positive shock from underlying economic growth rates, expected dividend growth increases. In response, the agent buys more of the stock that pays future dividends which increases its prices relative to dividends. This increases the quantity of risk that the agent bears. Since the market price of risk remains unchanged, overall equity-premia rises. Therefore, in response to a good shock, dividend yield decreases and expected return increases. From an equilibrium predictability point of view that is only feasible if the coefficient on dividend yield goes in the opposite direction from dividend yield from a shock in the growth rate. Let's re-write the expression for expected return in the form of a predictability relationship as

$$\mu_t^R = \left[\frac{(\gamma-1)\lambda}{\kappa+\beta}\sigma_x^2 G_X\right]\frac{D}{P}_t \tag{12}$$

where $\frac{D_t}{P_t} = \frac{1}{G(X_t)}$. We just established that the left hand side of this expression increases in X_t, and the dividend yield on the right decreases in X_t. However, the stochastic component

[2]Here, $\exp(\cdot) = \exp(P_1(\tau)X_t + P_2(\tau))$ is given in the appendix, and the expression above is positive due to a direct application of Cauchy-Schwartz inequality to functions $P_1(\tau)\sqrt{\exp(\cdot)}$ and $\sqrt{\exp(\cdot)}$, both of which are integrable in the domain as long as the transversality condition is satisfied.

of the coefficient on dividend yield, G_X, has the property that $\left(\frac{G_X}{G}\right)_X > 0$. This ensures that as $G(X_t)$ increases (dividend yield decreases), G_X also increases which "pulls up" a diminishing dividend yield to produce higher expected return. It is shown pictorially in Figure 3. This makes the coefficient of predictability itself time-varying - a fact empirically uncovered in Dangl and Thomas (2011) and also shown in Lettau and van Nieuwerburgh (2008). This stochastic nature of the predictability coefficient is missing in the equilibrium literature. It helps us understand how an economy can have simultaneously both high prices and high expected returns. It helps make returns stationary - when dividend yield changes, the return predicting coefficient moves in the opposite direction to keep returns stationary. The time-series property of the return predicting coefficient is magnified in the long-horizon as I show below.

The overall result suggests that as growth rate increases, expected dividend growth increases, the PD ratio increases (dividend yield decreases) and expected return increases. Thus, return shocks and dividend yield shocks are strongly negative correlated. Now I explore the effect of the time-variation in predictability coefficient for long-horizon returns.

1.3.1 Long Horizon Predictability

An investor with a long horizon holding period will invest P_t in the market at time t, and hold it until time T when the price will grow to P_T and he will also receive dividends from time t to T. Thus, his total return is given by

$$\bar{R}_T = \frac{P_T + \int_t^T D_r dr}{P_t} \tag{13}$$

This is a particular convenient way to pose the long-horizon predictability relationship because it is easier to solve. Notice I make one simplification where dividends are simply accumulated and not ploughed back into the stock. Note that this is different from the instantaneous excess return dynamics developed in (1.2.2), where I use $dR_t = \frac{dP_t + D_t dt}{P_t} - r_t^f dt = -E_t\left(\frac{dP}{P}, \frac{d\Lambda}{\Lambda}\right) + \cdots dW$. The latter expression, integrated forward to produce R_T, is wholly unsuitable in analyzing long-horizon cummulative returns. This is because the instantaneous cummulative return dynamics is of a dt - period return from t to $t + dt$ with dividends D_t and risk-free rate r_t^f held constant at time t. The expected growth rate X_t also stays constant, and I can only account for price change due to X_t. Integrating forward this quantity will not address the fact that there are dynamic relationships between prices, dividends and risk-free rate through the growth rate which will grow over time in longer horizon. To overcome this problem, I resort to looking at long horizon returns through the quantity in (13) where I accumulate dividends from t to T and also consider the intermediate shocks from dividend growth and risk-free rate to prices within holding period $z = T - t$.

First, I determine the dynamics of prices P_t. The full distribution of price growth from t to T can be written as

$$P_T = P_t \exp\left(\int_t^T \left[\mu_P(X_s) - \frac{1}{2}\sigma_P(X_s)\sigma_P(X_s)'\right] ds + \int_t^T \sigma_P(X_s) \cdot dW\right) \quad (14)$$

where $dW = [dW_D \;\; dW_x]$. The expressions for $\mu_P(X_t)$ and $\sigma_P(X_t)$ are in the appendix. $\mu_P(X_t)$ is the total change in price resulting from dividend growth, risk-free rate and compensation for bearing risk X_t along with other higher order terms whose effect over the long horizon could be substantial. $\sigma_P(X_t)$ is a vector of volatility shocks arising from both transient dividend shock and growth rate shock from X_t. They are determined by applying

Ito's Lemma to (8) and integrated forward. Similarly, dividend growth can be written as

$$D_r = D_t \exp\left(\int_t^r \left(X_s - \frac{1}{2}\sigma_D^2\right) ds + \int_t^r \sigma_D dW_D\right) \tag{15}$$

Substituting them both into (13), I can write total return from t to T as

$$\bar{R}_T = \left[G(X_t)\exp\left[\int_t^T \left[\mu_P(X_s) - \frac{1}{2}\sigma_P(X_s)\sigma_P(X_s)'\right] ds + \int_t^T \sigma_P(X_s)\cdot dW_s\right] + \int_t^T \exp\left[\int_t^r [X_s - \sigma_D^2]ds + \int_t^r \sigma_D dW_D\right] dr\right] \frac{D_t}{P_t} \tag{16}$$

This expresses cumulative return over horizon $z = T - t$ as a function of current dividend and growth rate shocks, as well as the effect of the entire path of the growth rates over the horizon. The first term inside the parenthesis is total price growth from t to T and the second term is the growth in dividends. At each point on the growth path, the risk-free rate and dividend growth expectation changes because of stochastic growth rate X_t leading to a change in price and, by extension, cumulative returns. It is straight-forward to see that the non-linearity of endogenous shocks rule out any possibility that the statistical properties of OLS will do justice in estimating the above expression.

Fortunately, the conditional expectation of the above expression in (16), has a more tractable form without the Brownian shocks. First, observe that the conditional expectation of future dividends has a closed-form solution.

Lemma 1.3.1 *Conditional expectation of future dividend satisfies*

$$E_t[D_r] = D_t \exp(A(s)X_t + B(s))$$

where $s = r - t$ and $A(s)$ and $B(s)$ are in the appendix.

Now it is straightforward to establish conditional expectation of cumulative expected return $\bar{R}_T$ using the accounting identity (13).

Proposition 1.3.1 *The price process in (14) implies $E_t\left[P_T\right] = P_t H(X_t, z)$ where $H(X_t, z) = E_t\left[\exp\left[\int_t^T \mu_P(X_s)ds\right]\right]$ with $z = T - t$. Then, using Lemma (1.3.1)*

$$
\begin{aligned}
E_t[\bar{R}_T] &= \frac{E_t[P_T] + \int_t^T E_t[D_r]ds}{P_t} \\
&= \left[G(X_t)H(X_t, z) + \int_t^T [\exp(A(r-t)X_t + B(r-t))]dr\right]\frac{D_t}{P_t} \\
&= \alpha(X_t, z)\frac{D_t}{P_t} \qquad (17)
\end{aligned}
$$

To convert the conditional expectation relationship into a percentage return form, I simply subtract one and focus on the quantity.

$$
E_t[\bar{R}_T] - 1 = \left[G(X_t)(H(X_t, z) - 1) + \int_t^T [\exp(A(r-t)X_t + B(r-t))]dr\right]\frac{D_t}{P_t} \qquad (18)
$$

The expression $H(X_t, z)$ is conditionally known at time t and represents expected price changes over horizon z due to the stochastic growth rate. It satisfies a second order partial differential equation that depends on X_t and z. It has an unique solution given a set of boundary condtions. One of them is natural $H(X_t, 0) = 1$ such that $\lim_{T\to t} E_t[P_T] = P_t$. However, due to all the non-linearties in X_t, its general form cannot be solved analytically, and no sensible boundary conditions are available in the X_t-plane to solve it numerically. Details are in the appendix. Thus, I resort to solving $H(X_t, z)$ by simulating several thousand paths of $X_{t\to T}$ to compute $E_t\left[\exp\left(\int_t^T \mu_P(X_s)ds\right)\right]$ for every initial point X_t.

The z-horizon return predictability coefficient $\alpha(X_t, z)$ is composed of two parts. There is an expected dividend growth component and then an expected price growth component as a dynamic response to dividend growth rates and dividend shocks. In traditional predictability regressions of Shiller (1981) and Fama-French (1988), the above conditional expectation relationship is tested by running univariate regression of cumulative returns of varying horizon on current dividend yields. The coefficients from these regressions are taken as constants and tests on the coefficients are performed using standard assymptotics. The structural relationship here suggests that the slope coefficient on these long-horizon regressions are themselves stochastic with crucial time-series properties, and as such, treating them as constants would lead to immense biases. The slope itself is a non-linear function of the underlying state variable that also affects the regressor and as such should contribute to the overall variance of the slope coefficient that treating it as a constant would miss. In fact, looking at the immense non-linearity of (16) in the Brownian shocks, it looks like the coefficient estimated via OLS will also be highly inconsistent. In fact, it confirms Valkanov's (2003) argument that the coefficient is a function of underlying shocks with fundamentally different properties than standard assymptotics which he analyzes by using the Functional Central Limit Theorem.

Another important aspect of these regressions is the explanatory power of the regression typically measured in terms of higher R^2-s as horizon increases. Fama and French (1998), for example, find R^2-s that range from 19% to as high as 64% over 1-5 year horizons. The equilibrium models of Bansal and Yaron (2004) and Campbell and Cochrane (1999) both show that return R^2-s are also increasing over the horizon. However, Goetzmann and Jorion (1993) and recent work of BRW (2008) have cast doubts on these findings. In the latter work, for example, the authors find that the R^2-s are not increasing but scale with time and are, in fact, decreasing slightly as return horizon increases. Goetzmann and Jorion (1993) show that one can still get high R^2-s and significant coefficients where there is no linear relationship between future returns and the dividend yield. The conditional mean relationship given in (17), provides a theoretical foundation to compute pseudo-R^2's in longer horizon. To gauge the magnitude of pseudo-R^2 from my structural model, I ask the question - *How much of the unconditional variance of $\bar{R}_T$ can be explained by the unconditional variance of the conditional mean relationship in (17)?* Thus, to infer the model implied R^2's for longer horizon, I simply compute pseudo-$R^2 = \frac{\mathrm{Var}\left(\alpha(X_t,z)\frac{1}{G(X_t)}\right)}{\mathrm{Var}(R_T)}$ where Var denotes unconditional variance. Notice that the coefficient of the conditional mean $\alpha(X_t, z)$ is itself a function of X_t which also impacts the dividend yield $\frac{1}{G(X_t)}$. Empirical works that treat the return predictability coefficient as a constant misses this extra uncertainty that increases with time.

It is also obvious from the expression of $\alpha(X_t, z = T - t)$ that two return predicting coefficients of different horizons T_1 and T_2 should also be correlated - not only through the current expected growth rate X_t, but also because they will share the same expected price and dividend growth changes upto $min(T_1, T_2)$. Naturally this persistance will be stronger if T_1 and T_2 are closer to each other than if they are further apart. Empirically, BRW (2008) has found that this correlation between the return predicting coefficients is quite significant. They are stronger when the horizons are closer as is the case in my model.

This establishes the full theory behind long horizon predictability that is completely endogenized within a one-channel Bansal and Yaron (2004) economy under DE preferences. The setting here is tractable enough to produce a semi closed-form estimate of the conditional mean of long-horizon regression with explicit expression for the long-horizon predictability coefficient. The result shows time-series dependence between the return predicting coefficient and dividend yield rendering inference drawn from pure OLS based exercises biased and inconsistent.

2 Empirical Methodology

2.1 A Bayesian Strategy

In order to get the parameter estimates that govern the above state-space, I follow a Bayesian methodology. Let the full parameter set that guides the system be $\theta = \{\mu_D, \mu_C, \sigma_D, \sigma_x, \kappa, \lambda\}$. The goal is to get joint estimates of $p(\theta, X)$ conditional on the data on consumption and dividend growth. Here, X denotes the full time-series of growth rates $\{X_1, \cdots, X_T\}$. We will follow a Markov Chain Monte Carlo (MCMC) algorithm that will draw them conditionally on each other

$$p(\theta|X) \qquad\qquad p(X|\theta)$$

In order to generate the parameters and the time-series of the growth rates, first I discretize dividend and consumption growth rates and write them in the familiar discrete-time state-space notation. Let g^d_{t+1} be dividend and g^c_{t+1} be consumption growth. Then the continuous-time state-space can be written by taking $dt = 1$ as

$$\begin{bmatrix} g^d_{t+1} \\ g^c_{t+1} \end{bmatrix} = \begin{bmatrix} (\mu_D + X_t) \\ (\mu_C + \lambda X_t) \end{bmatrix} + \begin{bmatrix} \sigma_D & 0 \\ 0 & \sigma_C \end{bmatrix} Z_1 \tag{19}$$

$$X_{t+1} = (1-\kappa)X_t + \sigma_x Z_2 \tag{20}$$

where $Z_1 \sim N(0, I_2)$ and $Z_2 \sim N(0, 1)$ are uncorrelated standard normals.

First, I draw the time-series of the growth rates X conditional on the rest of the parameter space, θ, and the full time-series of dividend and consumption growth. In order to draw the time-series of growth-rates, I follow a Bayesian version of kalman filter called

Forward Filtering Backward Sampling (FFBS) as introduced by Carter and Cohn (1996). In this step, recall I am assuming that I know the rest of the parameters θ, and I draw the full time-series of X given the full time-series of dividend and consumption growth.

Then, my goal is to draw the parameter set θ conditional on the full time-series of the growth rates, which I have obtained in the above step using FFBS. Here I generate the parameters using a MCMC algorithm called Gibbs sampler by which I draw one parameter at a time conditional on the rest of them - $\theta_i|\theta_{-i}, X, g^d, g^c$, where θ_{-i} is the rest of the parameters modulo the i-th one. In this simple state-space setting, all the posterior distributions of the parameters are available in elementary conjugate form. The exact parameters of these posterior distributions is discussed in detail in Allenby, McCulloch and Rossi (2005).

2.1.1 Priors

The strength of the Bayesian mechanism is the ability to specify prior information on the growth rate X since it is not directly observable. Prior belief on the parameters of X - κ and σ_x, based on the theory developed thus far allows incorporation of valuable economic intuition into the estimation process precisely because X_t is not directly observable. In order to generate high market prices of risk, I need the growth rate to be persistant. I impose a prior on $1-\kappa \sim N(0.95, 0.1^2)$. Furthermore, my choice of hyperparameters for prior on σ_x centers the prior mean of σ_x to be 0.015 - half the unconditional variance of aggregate consumption growth, with fairly high uncertainty which will show up in the posterior distribution of σ_x. Finally, since the theory heavily relies on $\lambda < 1$, I propose the prior $\lambda \sim N(0.40, 1^2)$. Note, these are all proper but extremely diffuse priors. The 95% confidence band of the prior distributions for these parameters is fairly wide and covers a broad range of possible values. I leave it up to the data to play a crucial role in identifying the posterior distribution of these parameters.

2.2 Data

I use US data from 1929-2010 sampled annually. Aggregate dividend data is from CRSP value-weighted portfolio. Cochrane (2008) points out that CRSP dividends capture all payments to investors - including cash mergers, liquidations and repurchases. The risk-free rate is obtained from the return on 90-day Treasury Bills. Aggregate consumption is non-durables and services divided through by population growth to make it per capita consumption. All nominal quantities are converted to real by deflating them by CPI.

3 Empirical Findings

The Gibbs sampler produces simulations of parameter values from their posterior distributions. The estimates from the state-space estimation (19)-(20) is reported in Table I in five different quintiles from 2.5-97.5-th quintile. It is clear from the MCMC simulated draws that the data has played a crucial role in pinning down the posterior distrubition of κ, σ_x and λ. The posterior distribution of both of these parameters have tightened around the posterior mean showing that the data provides valuable inference in mitigating the prior uncertainty about these parameters. The parameter for which the data plays the most crucial role is σ_x whose posterior mean is 0.027. Also, the time-series of growth rates that are filtered from aggregate consumption and dividend growth match the time-series behavior of the underlying series quite well as is shown in Figure 2.

Another important test whether the model parameters are meaningful is their ability to produce key moments of the macro data. The model implies that the unconditional mean, standard deviation and first order autocovariance of consumption growth are μ_C, $\sqrt{\lambda^2 \frac{\sigma_X^2}{1-(1-\kappa)^2} + \sigma_C^2}$ and $\lambda^2(1-\kappa)\frac{\sigma_X^2}{1-(1-\kappa)^2}$. Similarly, I can compute the unconditional moments of dividend growth. Table IB reports the posterior distribution of these moments computed from the posterior distribution of the parameters simulated via MCMC. The posterior distribution of the model implied moments match up very well with the data. The only statistic it falls short on is the correlation of dividend and consumption growth. Whereas in the data the correlation is 0.58, the model implied correlation is only between 0.30-0.47. That is primarily due to the fact that this a one factor model. With additional latent shocks, like stochastic volatility, this shortcoming can be easily addressed.

The overall message is that the parameters drawn from the MCMC can reproduce salient features of the macroeconomic data - the filtered draws of the states X_t can track the observed time-series of consumption and dividend growth and the parameter draws can match the key moments implied by the model.

3.1 Market Prices of Risk

To focus on asset pricing, I pick the following preference parameters - time discount parameter $\beta = .001$ and risk-aversion $\gamma = 7.5$. The posterior estimates of the market prices of risk is in Table II. There are two sources of risk in my economy - transient consumption volatility risk given by $\gamma\sigma_C$ and long-run risk from persistant growth rates given by $\frac{(\gamma-1)\lambda\sigma_x}{\kappa+\beta}$. The posterior distribution of the price of long-run risk dominates the price of transient volatility risk by a huge margin. Whereas the posterior mean of the price of transient risk is 0.15, the posterior mean of the price of long-run risk is 0.58. Clearly, the time-series of dividend and consumption risk implies that the magnitude of long-run risk is extremely economically significant. Hence, an agent in this economy with DE preferences is far more averse to marginal utility shocks resulting from long-run risk than from traditional transient consumption volatility shocks.

3.2 Asset Pricing

This subsection shows the quantative magnitude of key asset pricing quantities implied by my model. Taking the posterior distribution of the parameters, I simulate the posterior distribution of six key asset pricing quantities - expected excess return (9), volatility of cumulative return (10), dividend-price ratio (8), volatility of changes in dividend-price ratio, risk-free rate (7), the volatility of risk-free rate and the Sharpe Ratio. Since all of these quantities depend on the growth rate X_t, I integrate it out by using the stationary distribution of $X_t \sim N\left(0, \frac{\sigma_x}{\sqrt{2\kappa}}\right)$ to produce unconditional estimates. Table IV reports the 2.5-97.5-th quintiles of these quantities.

The model can match the equity premia (posterior distribution is 4.78-7.91%), dividend yield (posterior distribution is 3.16-4.56%) and the low discount rate, β, helps to match the risk-free rate (posterior distribution is 0.83-2.67%). The Appendix shows that for $\psi > 1$, I can generate a far lower risk-free rate with higher discount rates. At the same time, my model can also generate high volatilities of equity returns (posterior distribution is 16.22-19.37%) and risk-free rate (posterior distribution is 1.68-2.51%). To gauge the effect of long-run risk on equity volatility, notice that the volatility of changes in the dividend yield is between 9.14-12.37% which is solely determined by exposure to long-run risk.

The parameters of the state-space that match the time-series properties of consumption and dividend growth can generate plausible asset pricing quantities. Interestingly, there is only one source of priced risk and that is sufficient to generate all of these quantities. Clearly, additional factors, like stochastic volatility, can be used to enhance the quantitative effects. However, stochastic volatility is used in long-run risk models like Bansal and Yaron (2004) to generate time-variation in expected returns. That is not a requirement in my model. In fact, the time-variation in expected returns in this one-factor model generates the central result of this paper which is the time-variation in the predictability coefficient. That is discussed next.

3.3 Long Horizon Predictability

The standard iid view of the world implies that returns are unpredictable. However, standard regression tests like Campbell and Shiller (1988) show that high dividend-price (or earnings-price) ratios are correlated with high expected returns. Fama and French (1988) find significance in long-horizon return predictability. In terms of structural models, Campbell and Cochrane (1999) and Bansal and Yaron (2004) show through simulations that the long horizon coefficients are increasing in size over the horizon, are highly significant and the predictive power based on R^2-s from these regressions are also increasing. However, in their regression tests they treat the coefficient as a constant, whereas in my model it has its own time-varying properties.

In my case the long-horizon coefficient can be solved in a semi-closed form setting and is shown to be time-varying which has ramifications with regards to long-horizon predictability. Furthermore, I can isolate the effect of price growth and dividend growth on the coefficient. Both of these quantities are conditionally known and I can compute them without running any regressions. I also compute how informative this conditional relationship through a pseudo-R^2 ratio that measures the size of the unconditional variance of the conditional mean given by the predictability relationship relative to the unconditional variance of long-horizon return in my model.

Before I go on to long-horizon predictability, let us focus on the stochastic nature of the predictability coefficient.

$$\mu_t^R = G_X \frac{\lambda(\gamma - 1)\sigma_x^2}{\kappa + \beta} \frac{D_t}{P_t}$$

where $\frac{D_t}{P_t} = \frac{1}{G(X_t)}$. Taking the median values of parameters from Table I, I compute expected return, dividend yields and the stochastic component of the predictability coefficient G_X across high and low growth rates X_t that span the unconditional distribution of $X_t \sim N\left(0, \frac{\sigma_x}{\sqrt{2\kappa}}\right)$. As X_t increases, clearly expected dividend growth increases. In response, the bottom plot in Figure 3 shows that PD ratio increases, or dividend yield decreases. However, expected return in the top plot also increases. The only way that is feasible is if the return predicting coefficient increases and goes in the opposite direction from dividend yield. The middle graph of G_X suggests precisely that. It shows that the return predicting coefficient "pulls up"("pushes down") a decreasing(increasing) dividend yield to make expected return increasing (decreasing) in the growth rates. In essence, the return predicting coefficient shows the effect of risk-premia. When PD ratio is high due to high growth rates, the agent infers this "momentum" will continue because of positively autocorrelated growth rate shocks and buys more of the risky security. This increases the quantity of risk that he bears which increases the risk-premium. The opposite happens with decreasing growth rate shocks when the investor holds less of the risky asset thereby reducing the quantity of risk. The time-variation in the desire to bear risks is embodied in the time-variation in the predictability coefficient. Dangl and Halling (2011) documents precisely this time-variation in the predictability coefficient, and my model suggests a long-run risk based explanation for this phenomenon.

The z-horizon conditional predictability coefficient is previously shown to be

$$\left[G(X_t)(H(X_t, z) - 1) + \int_t^T [\exp(A(s)X_t + B(s))]ds\right]$$

that depends on the current growth rate X_t and the horizon $z = T - t$. The first component of the coefficient is expected price growth from $t \to T$, and the second reflects expected dividend growth. Taking the 2.5, 50 and 97.5-th quintiles of parameters and states drawn from the Gibbs sampler, I compute dividend growth and price growth at each point in time for time horizons 1, 3 and 5 years. The time-series of expected dividend growth is shown in Figures 5-7 and that of expected price growth in Figures 8-10.

First, let me focus on dividend growth. Assuming \$1 of dividends at time t, the graph shows expected dividend growth over horizons 1 (Figure 5), 3 (Figure 6) and 5 (Figure 7) years. The time-variation in expected dividend growth is substantial. In some periods, responding to poor negative shocks in the growth rate of dividends, the expected dividend growth falls below \$1. While the time-series is more stable in the post-war years, the early part of the sample around the Great Depression shows pronounced movements in expected dividend growth. Following the stock market crash in 1929, expected dividend growth fell precipitously. Then an increase in dividends, which came alongside a rebound in the stock market in the mid 1930's, shows large upswings in expected dividend growth which fell again when the stock market crashed in 1937. Subsequent boom during World War II lifted expected dividend growth, and the post-war time-series shows a lot less variability until the stock market crash of 2009. The time-series pattern is the same across all horizons although the magnitude of expected dividend growth changes substantially. The time-series average of expected dividend growth for each horizon is {1.01,1.04,1.08}, although there is a quite a bit of uncertainty behind those growth figures. The 2.5 and 97.5 quintiles for each of those quantities is {0.96,0.89,0.84} and {1.07,1.23,1.41}. Clearly, expected dividend growth rises over the horizon. But, the uncertainty also increases. For a 5-year horizon, on average, starting with \$1 in dividends expected dividend growth is anywhere from \$0.84 to \$1.41! This time-series variability in expected dividend growth is a lot more pronounced in price growth which is discussed next.

To give an intuition of what this term looks like, Figure 4 plots the $H(X_t, z) - 1$ function for different values of X_t. Expected price growth over any horizon is clearly monotonic in X_t, but it can be both positive and negative. Since growth rates are autocorrelated, if a negative shock is realized expected dividend growth rate will be negative for quite some time. Consequently, price of the asset which pays those dividends will also fall. Figures 8-10 show expected increases in PD ratio for horizons 1, 3 and 5 years, respectively. The time-series variation in expected price growth is a lot more pronounced than expected dividend growth primarily due to the leverage effect created by $\lambda < 1$. Whereas the post-war expected price growth is relatively stable, the largest increase takes place after the stock market rebounded in the mid-1930's. Following the market rebound, the market expectation of prices soared which dropped again when the stock market crashed in 1937. The World War II years saw more fluctuation in expectation of prices which stabilized at the end of the war. In the post-war years, the variability persists but not as pronounced as the pre-war years. As is true for expected dividend growth, expected price growth increases in horizon with substantial uncertainty. The time-series averages of price growth for horizons 1,3 and 5 years are {1.83, 3.60, 5.57}. However, the 2.5 and 97.5 quintiles are {0.45,-0.37,-0.90} and {4.35,11.77,20.20}. Clearly, the leverage effect exacerbates the effect of dividend growth rate shocks in prices. Interestingly, the price growth scales with the horizon. Dividing the median estimates by the respective horizons, median price growths become $\{1.83,1.20,1.11\}$. Since price growth clearly occupies the lion's share of the long-horizon predictability coefficient, the latter also scales with time - a fact uncovered empirically by BRW (2008).

Having shown time-series variation in the predictability coefficient, it is interesting to investigate what kind of unconditional claims we can make from this theory. Long-horizon predictability is handled in the equilibrium asset pricing literature by simulating long-horizons returns from the equilibrium model and running reduced form regressions. For references, see Bansal and Yaron (2004) Table VI and Campbell and Cochrane (1999) Table 5. Monotonically decreasing coefficients (on PD ratio) and increasing R^2's across horizons are taken as theoretical justification of the classic pattern in the reduced form works of Campbell and Shiller (1988) and Fama and French (1988). I perform the same regressions in Table V. They also show monotonically increasing coefficients (on DP ratio) and R^2's across the horizons. However, in my case these regressions are misspecified because of the dependence between the coefficient of long-horizon predictability and the dividend yield. Instead of relying on the R^2s as evidence of long-horizon predictability, I compute pseudo-R^2s described in Section 1.3.1. These pseudo-R^2s measure how much of the variance of long-horizon return in my model can be explained by the long-horizon predictability relationship and are shown in Table IV. Here as well, the pseudo-R^2s increase over the horizon, but with a big caveat. Across the horizons, the variance of the long-horizon predictability relationship increases much faster relative to the unconditional variance of long-horizon returns. In other words, pseudo-R^2s increase simply because the variance of the predictability relationship increases faster than the variance of total return for each horizon. This is hardly an evidence for long-horizon predictability. In fact, it shows that over long-horizon the predictability relationship has a lot of uncertainty, which should give us pause in rendering any qualitative judgement on long-horizon predictability. This result is due to the fact that in my model the regressor (DP ratio) and the predictability coefficient are jointly determined which increases the variance of the predictability relationship relative to other models where the coefficient is constant.

4 Conclusion

This paper shows that if aggregate consumption and dividends share a single slow-moving shock in the dynamics of their growth rate, then that has very important ramifications for the PD ratio under recursive preferences. Simple Mertonian mechanics imply elegant non-linearities in the PD ratio which creates stochastic volatility in returns and implies time-varying equity premium. This is a key contribution of this paper since the extant long-run risk literature (Bansal and Yaron (2004)) rely on stochastic volatility to generate time-variation in equity premium. Moreover, this non-linearity in PD ratio creates time-variation in the coefficient of predictability - an unexplored fact in the equilibrium asset pricing literature although empirical works with time-varying coefficients are promising. Parameters that can match key properties of consumption and dividend dynamics as well as basic asset pricing quantities imply large time-variation in the coefficient of predictability across all return horizons. However, there is one major caveat. The same parameters that replicate key macro and asset pricing quantities seem to imply a large uncertainty in the conditional mean of long-horizon predictability rendering statements about existence of predictability unreliable.

References

[1] Abel, Andrew, 1999, Risk Premia and Term Premia in General Equilibrium, *Journal of Monetary Economics*, 43, 3-33.

[2] Ang, Andrew and Jun Liu, 2007, Risk, Return and Dividends, *Journal of Financial Economics*, 85:1, 1-38.

[3] Ang, Andrew and Geert Bekaert, 2007, Stock Return Predictability: Is it There?, *Review of Financial Studies*, 20(3), 651-707.

[4] Attanasio, Orazio P., and Guglielmo Weber, 1989, Intertemporal Substitution, risk aversion and the Euler equation for consumption, *Economic Journal*, 99, 59-73.

[5] Avramov, Doron and Satadru Hore, 2007, Momentum, Information Uncertainty, and Leverage - an Explanation Based on Recursive Preferences, *Working Paper*.

[6] Bansal, Ravi, and Amir Yaron, 2004, Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles, *Journal of Finance*, 59(4), 1481-1509.

[7] Boudukh, Jacob, Matthew Richardson and Robert F. Whitelaw, 2008, The Myth of Long Horizon Predictability, *Review of Financial Studies*, 21(4), 1577-1605..

[8] Campbell, John Y., 1999, Asset Prices, Consumption and the Business Cycle, in John B. Taylor and Michael Woodford, eds.: *Handbook of Macroeconomics*, Volume 1, (Elsevier Science, North-Holland, Amsterdam)

[9] Campbell, John Y., 2001, Why Long Horizons? A study of Power Against Persistant Alternatives, *Journal of Empirical Finance*, 8, 459-491.

[10] Campbell, John Y. and John H. Cochrane, 1999, By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior, *The Journal of Political Economy*, 2, 205-251.

[11] Campbell, John Y. and R.J. Shiller, 1988, The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors, *Review of Financial Studies*, 1, 195-228.

[12] Carter, C.K. and R. Kohn, 1996, Markov Chain Monte Carlo in conditionally Gaussian state space models, *Biometrika*, 83,3, 589-601.

[13] Cochrane, John H., 2005, Asset Pricing, *Princeton University Press*, Princeton, NJ.

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

[15] Dangl, Thomas and Michael Halling, 2011, Predictive Regressions with Time-Varying Coefficients, Forthcoming, *Journal of Financial Economics*

[16] Duffie, Darrell and Larry Epstein1992Stochastic Differential UtilityEconometrica60:2353-394

[17] Fama, Eugene F. and Kenneth R. French, 1988, Dividend Yields and Expected Stock Returns, *Journal of Financial Economics*, 22, 3-25.

[18] Goetzmann, W. and Paul Jorion, 1993, Testing the Dividend Power of Predictive Yields, *Journal of Finance*, 48(2), 663-679.

[19] Guvenen, Faith, 2001, Mismeasurement of the elasticity of intertemporal substitution: The role of limited stock market participation, Unpublished manuscript, University of Rochester.

[20] Hall, Robert E., 1988, Intertemporal Substitution in Consumption, *Journal of Political Economy*, 96:2, 339-357.

[21] Hansen, Lars P., John Heaton and Nan Li, 2008, Consumption Strikes back? Measuring Long-Run Risk, *Journal of Political Economy*, 116:2, 260-302..

[22] Kandel, Shmuel and Robert F. Stambaugh, 1991, Asset Returns and Intertemporal Preferences, *Journal of Monetary Economics*, 27, 39-71.

[23] Lettau, Martin and Sydney C.Ludvigson, 2005, Expected Returns and Expected Dividend Growth, *Journal of Financial Economics*, 76, 583-626.

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

[25] Menzly, Lior, Tano Santos and Pietro Veronesi, 2004, Understanding Predictability, *Journal of Political Economy*, 112(1), 1-47.

[26] Rossi, Peter E., Greg M. Allenby and Robert McCulloch, 2005, *Bayesian Statistics and Marketing*, Wiley.

[27] Rozeff, Michael S., 1984, Dividend Yields are Equity Premiums, *Journal of Portfolio Management*, 11, 68-75.

[28] Schwert, G.W., 1989, Why Does Stock Market Volatility Change Over Time, *Journal of Finance*, 44(5), 1115-1153.

[29] Stambaugh, Robert F., 1999, Predictive Regressions, *Journal of Financial Economics*, 54, 375-421.

[30] Shiller, Robert J., 1984, Stock Prices and Social Dynamics, *Brookings Papers on Economic Activity*, 2, 457-498.

[31] Tallarini, Thomas D., 2000, Risk Sensitive Business Cycles, *Journal of Monetary Economics*, 45:3, 507-532.

[32] Valkanov, Ross, 2003, Long Horizon Regressions: Theoretical Results and Applications, *Journal of Financial Economics*, 68, 201-232.

[33] Vissing-Jorgensen, Annette, 2002, Limited Asset Market Participation and the Elasticity of Intertemporal Substitution, *Journal of Political Economy*, 110:4, 825-853.

[34] Yogo, Motohiro, 2004, Estimating the Elasticity of Intertemporal Substitution when Instruments are Weak, *Review of Economics and Statistics*, 2004, 86(3).797-810

[35] Weil, Philippe, 1989, The Equity Premium Puzzle and the Risk-Free Rate Puzzle, *Journal of Monetary Economics*, 24, 401-421.

5 Appendix

Proof of (1.1.1) The Bellman equation in (4) can be written as

$$J_C C[\mu_C + \lambda X_t] - J_X \kappa X_t + \frac{1}{2} J_{CC} C^2 \sigma_C^2 + \frac{1}{2} J_{XX} \sigma_X^2 + f(C, J) = 0$$

The continuation utility J has a solution of the form

$$(1-\gamma)J = \exp(u_0 \ln C_t + u_1 X_t + u_2)$$

Substituting it in and collecting terms, reduces the above equation to a system of ODE's that can be solved recursively

$$\begin{aligned} u_0 &= (1-\gamma) \\ u_1 &= \frac{(1-\gamma)\lambda}{\kappa + \beta} \\ u_2 &= \frac{(1-\gamma)}{\beta} \left[\mu_C - \frac{1}{2}\gamma\sigma_C^2 + \frac{\lambda^2 (1-\gamma)\sigma_x^2}{2(\kappa+\beta)^2} \right] \end{aligned}$$

Thus, the continuation utility function reduces to $J(C_t, X_t) = \frac{C_t^{1-\gamma}}{1-\gamma} \exp(u_1 X_t + u_2)$.

Proof of Proposition (1.2.1) The pricing kernel for stochastic differential utility can be written as

$$\frac{d\Lambda}{\Lambda} = \frac{df_C}{f_C} + f_J dt$$

Using the above utility function, let $g = f_C = \frac{\beta(1-\gamma)J}{C} = \beta C^{-\gamma} \exp(u_1 X_t + u_2)$ and $f_J = -\beta(1 + u_1 X + u_2)$. Use Ito's Lemma on g and (2) and (3) one can rewrite the pricing kernel as

$$
\begin{aligned}
\frac{d\Lambda}{\Lambda} &= -r_t^f dt - \gamma\sigma_C dW_C - \frac{\lambda(\gamma-1)}{\kappa+\beta}\sigma_x dW_X \\
r_t^f &= \lambda X_t + \mu_C - \gamma\sigma_C^2 + \beta
\end{aligned}
$$

Proof of Propostion (1.2.2**)** The stock price of a firm is

$$
\begin{aligned}
P_t &= \frac{1}{\Lambda_t} E_t \int_t^\infty \Lambda_s D_s ds \\
&= \frac{1}{\Lambda_t} \int_t^\infty E_t \Lambda_s D_s ds
\end{aligned}
$$

Define $h_t = \Lambda_t D_t$. Thus

$$
\frac{dh}{h} = [(1-\lambda)X_t + \mu_D - \mu_C + \gamma\sigma_C^2 - \beta]dt - \gamma\sigma_C dW_c - \frac{\lambda(\gamma-1)\sigma_x}{\kappa+\beta} dW_x + \sigma_D dW_D
$$

Applying Feynman-Kac, $E_t[\Lambda_s D_s] = f(\Lambda_t D_t, X_t, s-t) = f(h_t, X_t, \tau = s-t)$. Applying Ito's Lemma to f and the martingale restriction, I get the following PDE

$$
f_h h[(1-\lambda)X_t + \mu_D - \mu_C + \gamma\sigma_C^2 - \beta] - f_X \kappa X_t + \frac{1}{2}\left(f_{hh} dh^2 + f_{XX}\sigma_x^2\right) - f_{hX}\frac{\lambda(\gamma-1)\sigma_x^2}{\kappa+\beta} - f_\tau = 0
$$

Guess a solution of the form $f = h_t \exp(P_1(\tau)X_t + P_2(\tau))$. Plug the solution in the above PDE and after collecting the terms in the constant and X_t, I get a system of ODE's of the form

$$
\begin{aligned}
P_1'(\tau) &= (1-\lambda) - \kappa P_1(\tau) \\
P_2'(\tau) &= \mu_D - \mu_C + \gamma\sigma_C^2 - \beta - P_1(\tau)\sigma_x^2\left[\frac{\lambda(\gamma-1)}{\kappa+\beta} - \frac{1}{2}P_1(\tau)\right]
\end{aligned}
$$

with initial conditions $P_1(0) = P_2(0) = 0$. The solution of these ODEs are

$$\begin{aligned}
P_1(\tau) &= \frac{1-\lambda}{\kappa}\left(1 - e^{-\kappa\tau}\right) \\
P_2(\tau) &= a\tau + b(e^{-\kappa\tau} - 1) + c(1 - e^{-2\kappa\tau}) \\
a &= \mu_D - \mu_C + \gamma\sigma_C^2 - \beta + \frac{\sigma_x^2(1-\lambda)}{2\kappa}\left[\frac{1-\lambda}{\kappa} - 2\frac{\lambda(\gamma-1)}{\kappa+\beta}\right] \\
b &= \frac{1-\lambda}{\kappa}\left[\frac{\sigma_x^2}{\kappa}\left[\frac{1-\lambda}{\kappa} - \frac{\lambda(\gamma-1)}{\kappa+\beta}\right]\right] \\
c &= \frac{\sigma_x^2(1-\lambda)^2}{4\kappa^3}
\end{aligned}$$

Thus, $E_t\left[\Lambda_s D_s\right] = \Lambda_t D_t \exp(P_1(\tau)X_t + P_2(\tau))$ which implies

$$P_t = D_t G(X_t)$$

where $G(X_t) = \int_t^\infty \exp(P_1(\tau)X_t + P_2(\tau))ds$. The transversality condition holds for $a < 0$.

Cumulative excess return $dR_t = \frac{D_t dt + dP - r_t^f dt}{P_t}$ over a small interval dt is

$$dR_t = \mu_t^R dt + \sigma_D dW_D + \frac{G_X}{G}\sigma_x dW_x$$

where $\mu_t^R = -Cov_t\left(\frac{d\Lambda_t}{\Lambda_t}, dP_t\right) = \frac{G_X}{G}\frac{\lambda(\gamma-1)}{\kappa+\beta}\sigma_x^2$.

Long Horizon Predictability: The expression for $z = T - t$-horizon total return is

$$\bar{R}_T = \frac{P_T + \int_t^T D_r dr}{P_t}$$

To compute the expression for long run predictability, first let us write down the SDE that

G satisfies:

$$dG = \mu_G dt + \sigma_G dW_x$$

where

$$\begin{aligned}
\mu_G &= \frac{\sigma_x^2}{2}\int_t^\infty \exp(\cdot)P_1^2(\tau)ds - \kappa X_t \int_t^\infty \exp(\cdot)P_1(\tau)ds \\
&\quad - \int_t^\infty \exp(\cdot)\left(P_1'(\tau)X_t + P_2'(\tau))\right)ds - 1 \\
\sigma_G &= \sigma_x \int_t^\infty \exp(\cdot)P_1(\tau)d\tau
\end{aligned}$$

Furthermore, since $P_t = D_t G(X_t)$, then

$$\begin{aligned}
\frac{dP}{P} &= \left[\mu_D + X_t + \frac{\mu_G}{G}\right]dt + \sigma_D dW_D + \frac{\sigma_G}{G}dW_x \\
&= \mu_P(X_t)dt + \sigma_P(X_t)\cdot dW \qquad (21)
\end{aligned}$$

where $\mu_P(X_t) = \left[\mu_D + X_t + \frac{\mu_G}{G}\right]$ and $\sigma_P(X_t) = \left[\sigma_D \quad \frac{\sigma_G}{G}\right]$ and $dW = [dW_D \quad dW_x]$. In integral form, that can be expressed as

$$P_T = P_t \exp\left[\int_t^T \left[\mu_P(X_s) - \frac{1}{2}\sigma_P(X_s)\sigma_P(X_s)'\right]ds + \int_t^T \sigma_P(X_s)\cdot dW_s\right] \qquad (22)$$

The dividend process in (1) can be written as $D_r = D_t \exp\left[\int_t^r [X_s - \frac{1}{2}\sigma_D^2]ds + \int_t^r \sigma_D dW_D\right]$.

Thus, z-horizon return can be written as

$$\bar{R}_T = \left[G(X_t)\exp\left[\int_t^T\left[\mu_P(X_s) - \frac{1}{2}\sigma_P(X_s)\sigma_P(X_s)'\right]ds + \int_t^T \sigma_P(X_s)\cdot dW_s\right] + \int_t^T \exp\left[\int_t^r [X_s - \sigma_D^2]ds + \int_t^r \sigma_D dW_D\right]dr\right]\frac{D}{P} \qquad (23)$$

Fortunately, the conditional expectation of $\bar{R}_T$ has an easier form. First, I need to compute $E_t[P_T] = f(P_t, X_t, z = T-t)$. Applying Feynman-Kac to f and enforcing the martingale restriction produces the PDE,

$$f_P P\left[\mu_D + X_t + \frac{\mu_G}{G}\right] - f_x \kappa X_t + \frac{\sigma_x^2}{2} f_{XX} + \frac{P^2}{2} f_{PP}\left(\sigma_D^2 + \frac{\sigma_G^2}{G^2}\right) - f_z + P f_{PX}\frac{\sigma_G \sigma_x}{G} = 0$$

Notice that the above PDE is homogeneous of degree 1 in P_t. Thus, I can propose a solution of the form $f = P_t H(X, z)$ which reduces it to

$$\left[\mu_D + X_t + \frac{\mu_G}{G}\right] - \frac{H_X}{H}\kappa X_t + \frac{\sigma_x^2}{2}\frac{H_{XX}}{H} + \frac{H_X}{H}\frac{\sigma_G \sigma_x}{G} = \frac{H_z}{H}$$

with boundary condition $H(X_t, 0) = 1$. Thus $E_t[P_T] = P_t H(X_t, z)$. Using (22) and the law of iterated expectations, I can write $E_t[P_T] = P_t E_t\left[\exp\left[\left(\int_t^T \mu_P(X_s)ds\right)\right]\right]$ which implies $H(X_t, z) = E_t\left[\exp\left(\int_t^T \mu_P(X_s)ds\right)\right]$ which satisfies the boundary condition that $H(X_t, 0) = 1$. The conditional expectation of dividends can be obtained from a direct application of Feynman-Kac and can be solved in closed-form. $E_t[D_r] = D_t \exp\left[A(s)X_t + B(s)\right]$ where

$$\begin{aligned} A(s;\kappa) &= \frac{1 - e^{-\kappa s}}{\kappa} \\ B(s) &= \mu_D s + \frac{\sigma_x^2}{2\kappa^2}\left(s - 2A(s;\kappa) + A(s;2\kappa)\right) \end{aligned}$$

and $s = r - t$. Now, conditional expectation of cumulative return over any horizon from T

to t can be written as

$$
\begin{aligned}
E_t[\bar{R}_T] &= \frac{E_t[P_T + \int_t^T D_r dr]}{P_t} \\
&= \frac{E_t[P_T] + \int_t^T E_t[D_r] dr}{P_t} \\
&= \left[G(X_t) H(X_t, z) + \int_t^T [\exp(A(r-t)X_t + B(r-t))] dr \right] \frac{D_t}{P_t} \\
&= \alpha(X_t; T, t) \frac{D_t}{P_t}
\end{aligned}
$$

$H(X_t; T, t) = E_t \left[\exp \left[\int_t^T \mu_P(X_s) ds \right] \right]$. Thus, the conditional mean of cumulative return depends on the whole path of the growth rates X_s from t to T which can be generated given an initial X_t.

Unfortunately, the conditional or unconditional variance has no easy formulation. One has to simulate the full sample path of $\bar{R}_T$ according to (23) and compute the variance based on simulation.

Price-Dividend Ratio for $\psi \neq 1$: The above analysis holds for $\psi = 1$. Here I show that for $\psi \neq 1$, the price-dividend ratio is isomorphic to the $\psi = 1$ case. Hence, the predictability results that I derived earlier would hold for $\psi \neq 1$ as well. More specifically, I show in this section that the positive relationship between growth rates and the price-dividend ratio - the centerpiece of our above analysis, holds here for $\psi > 1$.

The normalized aggregator for the general ψ case is given by

$$
f(C, J) = \frac{\beta(1-\gamma)J}{1 - \frac{1}{\psi}} \left[C^{1-\frac{1}{\psi}} \left((1-\gamma)J\right)^{\frac{\frac{1}{\psi}-1}{1-\gamma}} - 1 \right]
$$

The Bellman equation still takes the form

$$J_C C[\mu_C + \lambda X_t] - J_X \kappa X_t + \frac{1}{2} J_{CC} C^2 \sigma_C^2 + \frac{1}{2} J_{XX} \sigma_X^2 + f(C, J) = 0$$

where $f(C, J)$ now takes the general form. Guess a solution of the form $J = \frac{C^{1-\gamma}}{1-\gamma} g(X_t)$ and plug it into the Bellman equation. It reduces to

$$\frac{\beta}{1-\frac{1}{\psi}} \left[g^{\frac{\frac{1}{\psi}-1}{1-\gamma}} - 1 \right] + \mu_C + \lambda X_t - \frac{\gamma}{2} \sigma_C^2 - \frac{g_X}{g} \frac{\kappa}{1-\gamma} X_t + \frac{\sigma_X^2}{2(1-\gamma)} \frac{g_{XX}}{g} = 0 \tag{24}$$

The pricing kernel takes the form

$$\begin{aligned}
\frac{d\Lambda}{\Lambda} &= -r_t^f dt - \gamma \sigma_C dW_C - \frac{\gamma - \frac{1}{\psi}}{1-\gamma} \frac{g_X}{g} \sigma_X dW_X \\
r_t^f &= \beta + \frac{\mu_C}{\psi} - \frac{\gamma \left(1 + \frac{1}{\psi}\right) \sigma_C^2}{2} - \frac{\sigma_X^2}{2} \frac{\left(\gamma - \frac{1}{\psi}\right)\left(1 - \frac{1}{\psi}\right)}{(1-\gamma)^2} \left(\frac{g_X}{g}\right)^2 + \frac{\lambda}{\psi} X_t
\end{aligned}$$

In order to price assets, I need a solution of the function g which should satisfy the functional relationship given by (24).

First, I will solve for the price of discounted future consumption, and then look for a solution of g around the unconditional mean of the consumption-wealth ratio. The discounted price of future consumption is given by

$$W_t = \frac{1}{\Lambda_t} E_t \int_t^\infty \Lambda_s C_s ds$$

Applying Fubini's Theorem and taking standrad limits (refer to Cochrane(2005) Pages

27-29), the consumption wealth ratio is given by the relationship

$$\frac{C_t}{W_t}dt = r_t^f dt - E_t\left[\frac{dW}{W}\right] - E_t\left[\frac{d\Lambda}{\Lambda}\frac{dW}{W}\right] \tag{25}$$

Guess $W_t = C_t \frac{g(X_t)^{\frac{1-\frac{1}{\psi}}{1-\gamma}}}{\beta}$. Applying Ito's Lemma

$$\frac{dW}{W} = \left[\mu_C + \lambda X_t - \frac{1-\frac{1}{\psi}}{1-\gamma}\frac{g_X}{g}\kappa X_t + \frac{1-\frac{1}{\psi}}{1-\gamma}\frac{\sigma_X^2}{2}\left[\frac{g_{XX}}{g} + \frac{\gamma-\frac{1}{\psi}}{1-\gamma}\left(\frac{g_X}{g}\right)^2\right]\right]dt + \sigma_C dW_C + \frac{1-\frac{1}{\psi}}{1-\gamma}\frac{g_X}{g}\sigma_X dW_X$$

Plugging in wealth dynamics, risk-free rate and the pricing kernel into (25), I get

$$\begin{aligned}
\frac{C_t}{W_t} &= \beta + \left(\frac{1}{\psi} - 1\right)\left(\mu_C + \lambda X_t - \frac{\gamma\sigma_C^2}{2} - \frac{g_X}{g}\frac{\kappa}{1-\gamma}X_t + \frac{\sigma_X^2}{2(1-\gamma)}\frac{g_{XX}}{g}\right) \\
&= \beta + \left(\frac{1}{\psi} - 1\right)\left(1 - g^{\frac{\frac{1}{\psi}-1}{1-\gamma}}\right)\frac{\beta}{1-\frac{1}{\psi}} \\
&= \beta g^{\frac{\frac{1}{\psi}-1}{1-\gamma}}
\end{aligned}$$

The second line follows from the first line due to the Bellman equation restriction in (24). This confirms that my choice of consumption-wealth ratio is right. In fact, as $\psi \to 1$, the consumption-wealth ratio approaches β which is a familiar result for unit elasticity of intertemporal substitution.

Now, let $\mu = E\left[\ln \frac{C}{W}\right]$. A first-order approximation of the consumption to wealth ratio around μ produces

$$\beta g^{\frac{\frac{1}{\psi}-1}{1-\gamma}} = \frac{C_t}{W_t} \approx e^{\mu}(1-\mu) + e^{\mu}\left(\ln\beta + \frac{\frac{1}{\psi}-1}{1-\gamma}\ln g\right)$$

Substituting the approximation above into (24), the original Bellman equation reduces to

$$\frac{1}{1-\frac{1}{\psi}}\left[e^{\mu}(1-\mu)+e^{\mu}\left(\ln\beta+\frac{\frac{1}{\psi}-1}{1-\gamma}\ln g\right)-\beta\right]+\mu_C+\lambda X_t-\frac{\gamma}{2}\sigma_C^2-\frac{g_X}{g}\frac{\kappa}{1-\gamma}X_t+\frac{\sigma_X^2}{2(1-\gamma)}\frac{g_{XX}}{g}=0$$

This has the familiar exponentially affine solution $g(X_t)=e^{u_1X_t+u_2}$, where u_1 and u_2 are given by

$$\begin{aligned}
u_1 &= \frac{\lambda(1-\gamma)}{\kappa+e^{\mu}}\\
u_2 &= \frac{1-\gamma}{1-\frac{1}{\psi}}\left[1-\mu+\ln\beta-e^{-\mu}\beta\right]+\frac{1-\gamma}{e^{\mu}}\left[\mu_C-\frac{\gamma\sigma_C^2}{2}+\frac{\sigma_X^2\lambda^2(1-\gamma)}{2(\kappa+e^{\mu})^2}\right]
\end{aligned}$$

Now, the pricing kernel and risk-free takes the form of

$$\begin{aligned}
\frac{d\Lambda}{\Lambda} &= -r_t^f dt-\gamma\sigma_C dW_C-\frac{\left(\gamma-\frac{1}{\psi}\right)\lambda}{\kappa+e^{\mu}}\sigma_X dW_X\\
r_t^f &= \beta+\frac{\mu_C}{\psi}-\frac{\gamma\left(1+\frac{1}{\psi}\right)\sigma_C^2}{2}-\frac{\left(\gamma-\frac{1}{\psi}\right)\left(1-\frac{1}{\psi}\right)\sigma_X^2\lambda^2}{2\left(\kappa+e^{\mu}\right)^2}+\frac{\lambda}{\psi}X_t\\
&= A+BX_t
\end{aligned}$$

Notice that as $\psi\to 1$, the consumption to wealth ratio converges to β, i.e. $\mu\to\ln\beta$ as $\psi\to 1$. Plugging in that limit makes the function g, risk-free rate and risk-prices converge to their $\psi=1$ limit derived in the previous section. Thus, this method can also be considered to be an approximate solution around $\psi=1$.

At this point, I apply the same methodology as in the previous section to derive the price-dividend ratio which takes the form

$$G(X_t)=\int_t^{\infty}\exp(P_1(\tau)X_t+P_2(\tau)ds,$$

where $\tau = s - t$. $P_1(\tau)$ and $P_2(\tau)$ are given in closed form as

$$\begin{aligned}
P_1(\tau) &= \frac{1-\frac{\lambda}{\psi}}{\kappa}\left(1-e^{-\kappa\tau}\right) \\
P_2(\tau) &= a\tau + b\left(e^{-\kappa\tau}-1\right) + c\left(1-e^{-2\kappa\tau}\right) \\
a &= \mu_D - A + \frac{\sigma_X^2\left(1-\frac{\lambda}{\psi}\right)}{2\kappa}\left[\frac{1-\frac{\lambda}{\psi}}{\kappa} - 2\frac{\left(\gamma-\frac{1}{\psi}\right)\lambda}{\kappa+e^{\mu}}\right] \\
b &= \frac{\sigma_X^2\left(1-\frac{\lambda}{\psi}\right)}{\kappa^2}\left[\frac{1-\frac{\lambda}{\psi}}{\kappa} - \frac{\left(\gamma-\frac{1}{\psi}\right)\lambda}{\kappa+e^{\mu}}\right] \\
c &= \frac{1}{\kappa}\left(\frac{\sigma_X\left(1-\frac{\lambda}{\psi}\right)}{2\kappa}\right)^2
\end{aligned}$$

As $\psi \to 1$, P_1 and P_2 converge to the solutions derived in the earlier section. The risk-premia in this case is given by $\mu_t^R = \frac{G_X}{G}\frac{\lambda\left(\gamma-\frac{1}{\psi}\right)}{\kappa+e^{\mu}}\sigma_X^2$.

First of all, notice that for expected excess return to be positive, we need early resolution of uncertainty, i.e $\gamma > \frac{1}{\psi}$. The central predictability result derived in the earlier section depended on $\frac{G_X}{G} > 0$. For $\psi \neq 1$, this quantity will be positive as long as $1 - \frac{\lambda}{\psi} > 0$. We have estimated λ to be far less than one, and thus if $\psi > 1$, that ensures $\frac{G_X}{G} > 0$ for $\psi \neq 1$. Risk-premia was pro-cyclical in the previous section as long as $\gamma > 1$. In this case, risk-premia is pro-cyclical as long as $\gamma > \frac{1}{\psi}$ which holds if γ and ψ are both greater than one.

Thus, the predictability relationship derived in closed form for $\psi = 1$ in the previous section will also hold in the $\psi > 1$ setting.

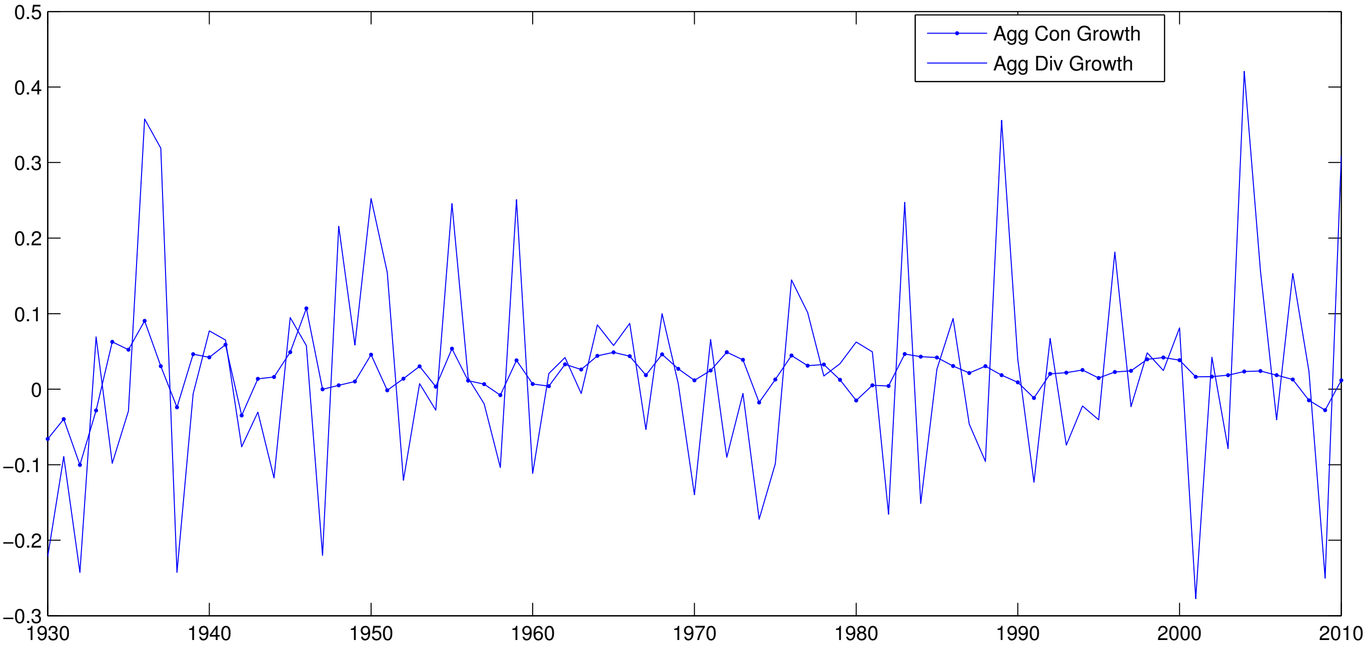

Figure 1: *Time-Series Plot of Aggregate Dividend Growth against Aggregate Consumption Growth.*

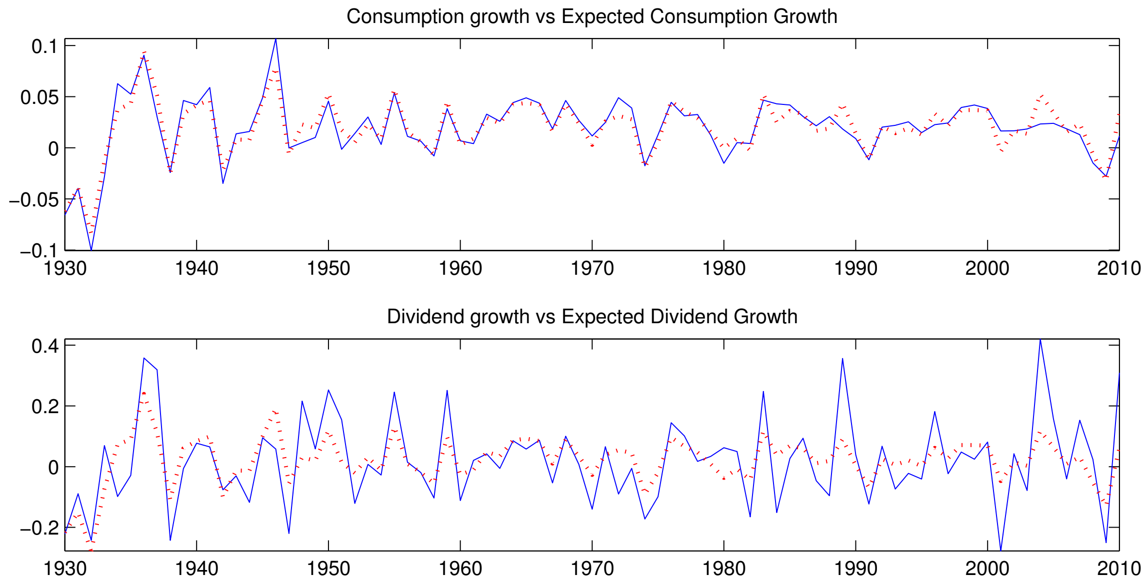

Figure 2: *The top graph shows median expected consumption growth rate (dotted line in red) against actual consumption growth (solid line in blue). From the posterior distribution of parameters and the latent state* X_t*, I form the posterior expected consumption growth rate using* $\mu_C + \lambda X_t$ *and report the median of the time-series. The bottom graph shows median expected dividend growth rate* $\mu_D + X_t$ *against dividend growth.*

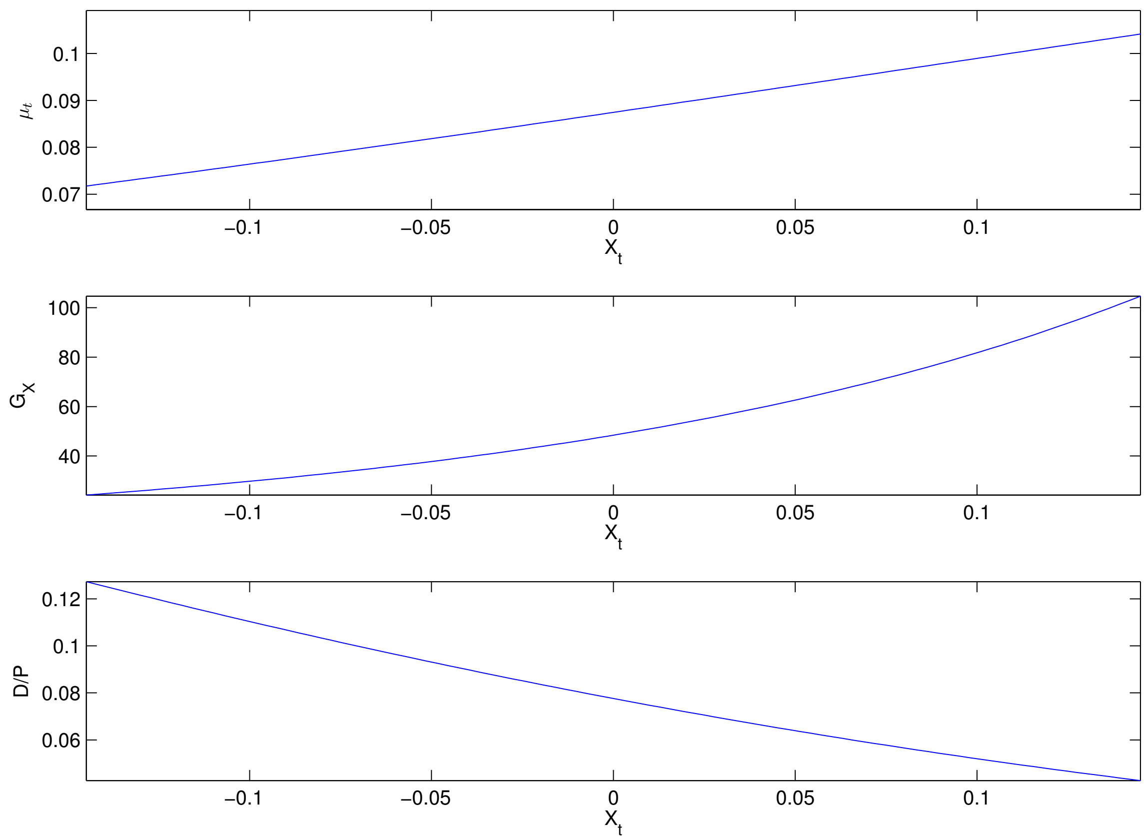

Figure 3: *The top panel plots expected excess return (9) across different growth rates, X_t. The middle panel plots G_X and the bottom panel plots the dividend price ratio $\frac{1}{G}$. The parameters used are the median parameters which are summarized in Table I.*

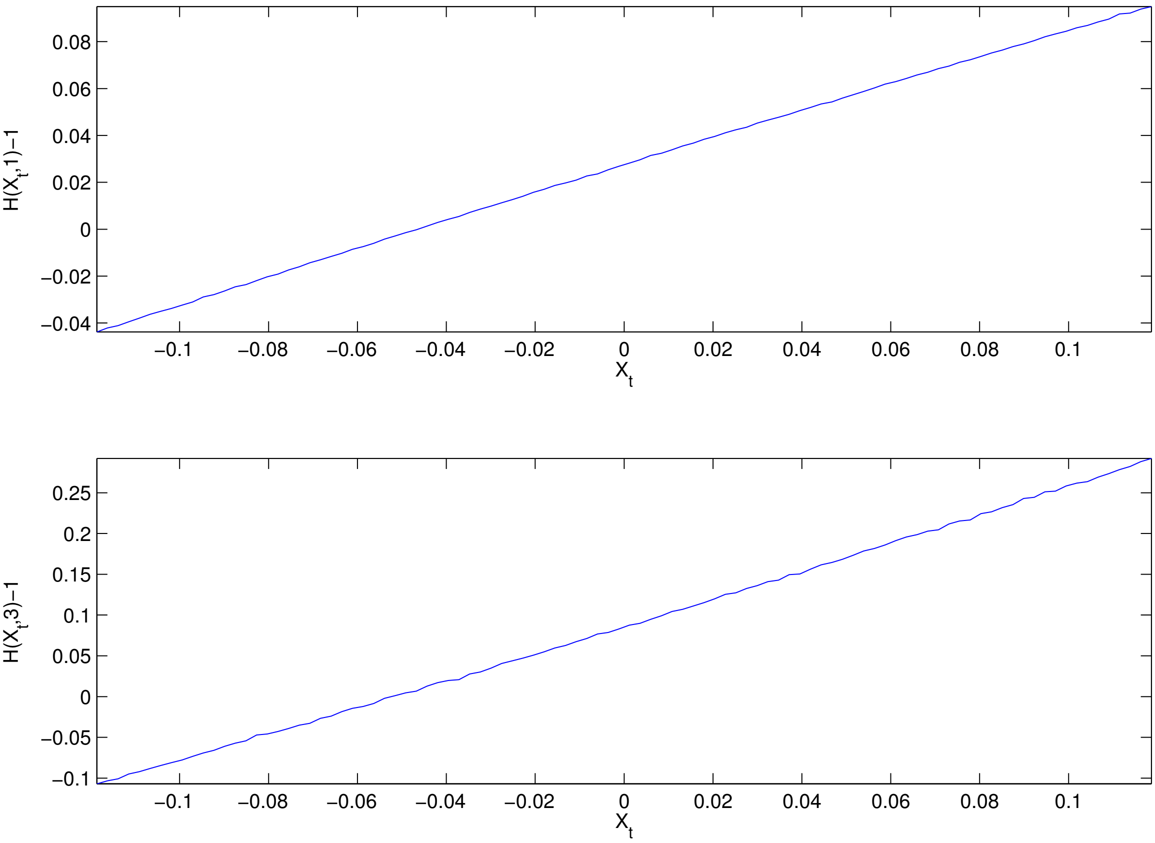

Figure 4: *The function* $H(X_t, z) - 1$ *for* $z = 1, 3$ *for different* $X's$ *taken from the unconditional distribution of* $X_t \sim N\left(0, \frac{\sigma_x}{\sqrt{(2\kappa)}}\right)$ *using the median parameter values in Table I.*

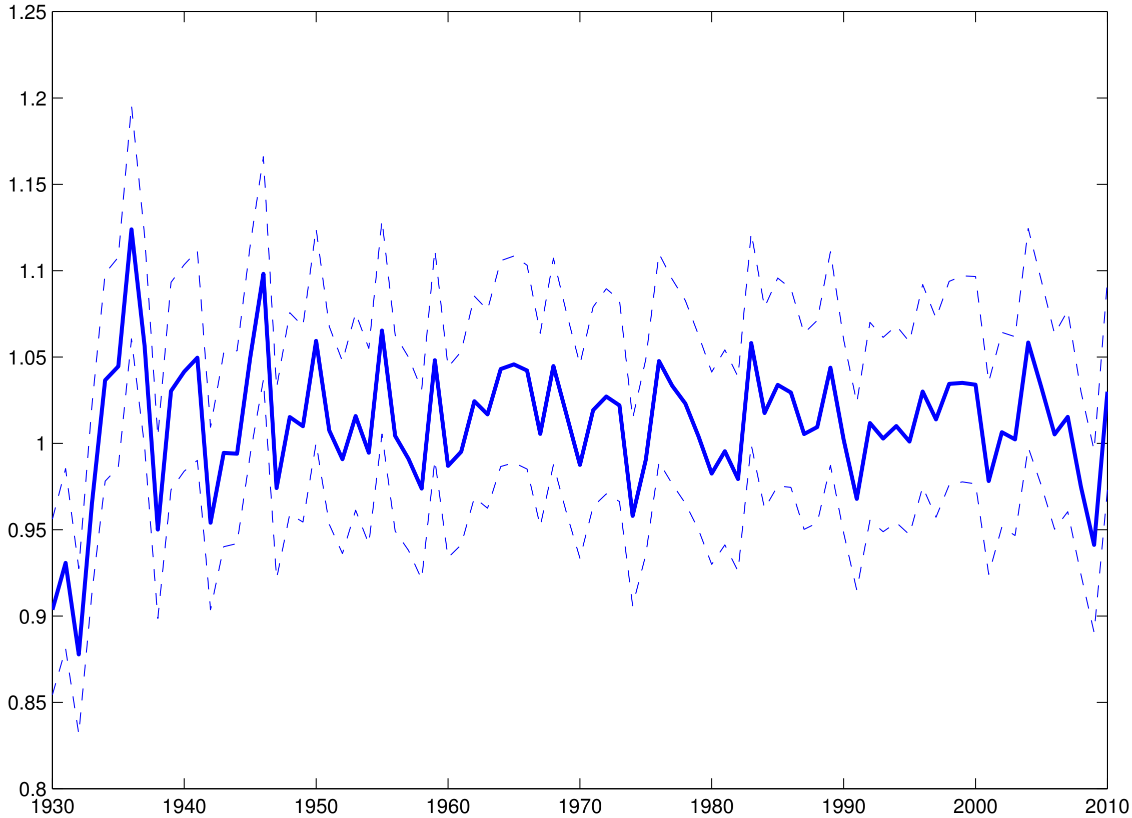

Figure 5: *Time-series of expected future dividends over a horizon of 1 year. Taking 2.5, 50 and 97.5th quintile of parameters and states* X_t *obtained from the MCMC, I plot expected future dividends in one year using Lemma 1.3.1.*

.

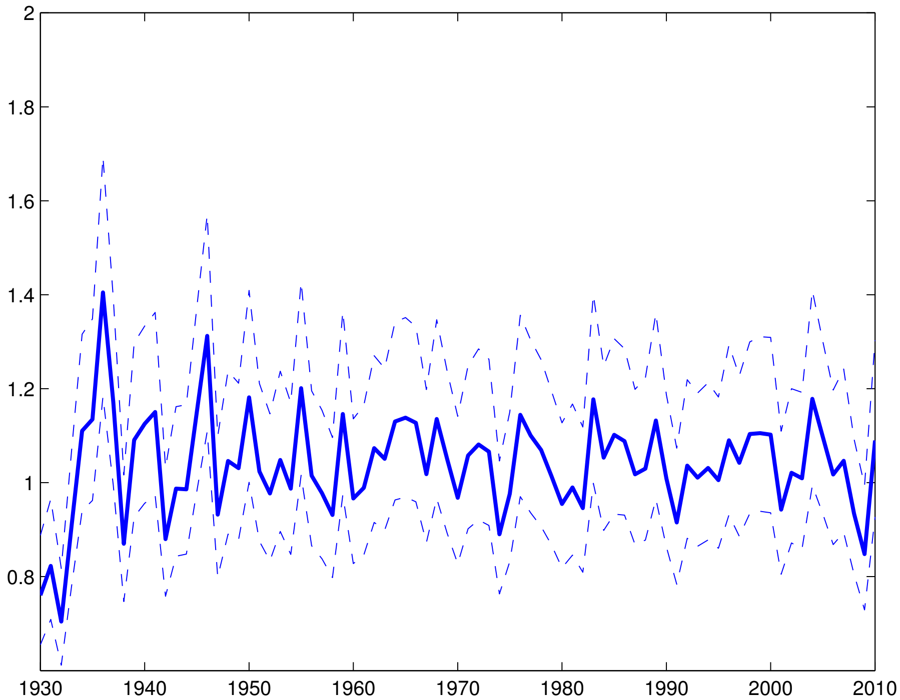

Figure 6: *Time-series of expected future dividends over a horizon of 3 years. Taking 2.5, 50 and 97.5th quintile of parameters and states X_t obtained from the MCMC, I plot expected future dividends in three years using Lemma 1.3.1.*

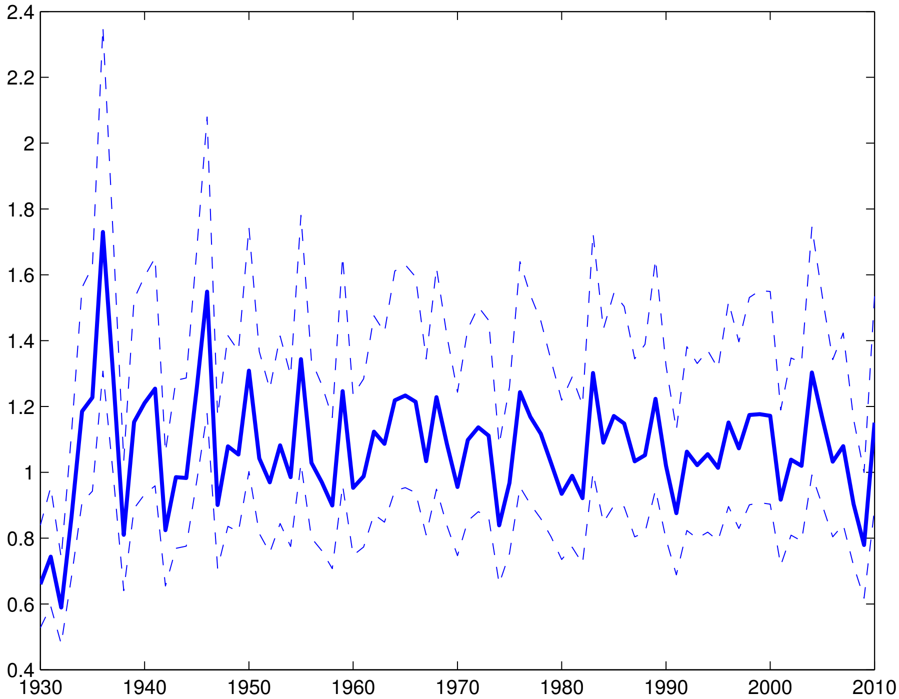

Figure 7: *Time-series of expected future dividends over a horizon of 5 years. Taking 2.5, 50 and 97.5th quintile of parameters and states X_t obtained from the MCMC, I plot expected future dividends in five years using Lemma 1.3.1.*

.

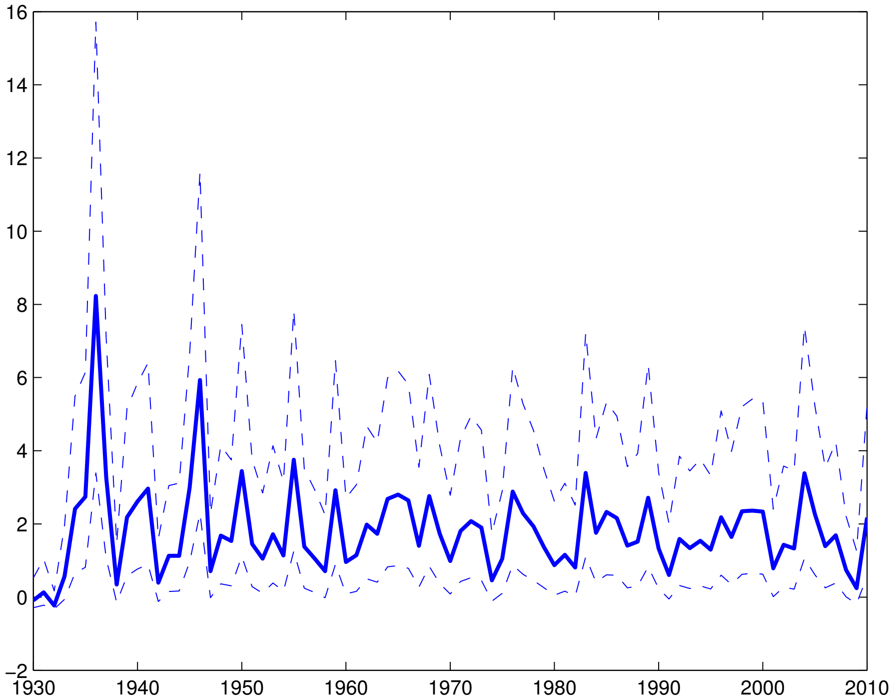

Figure 8: *Time-series of price-growth over a horizon of 1 year. Taking 2.5, 50 and 97.5th quintile of parameters and states* X_t *obtained from the MCMC, I plot expected price growth in one year using* $H(X_t, 1) = E_t\left[\exp\left(\int_t^1 \mu_P(X_s)ds\right)\right]$. *The expression for* $\mu_P(X_s)$ *is given in the appendix.*

.

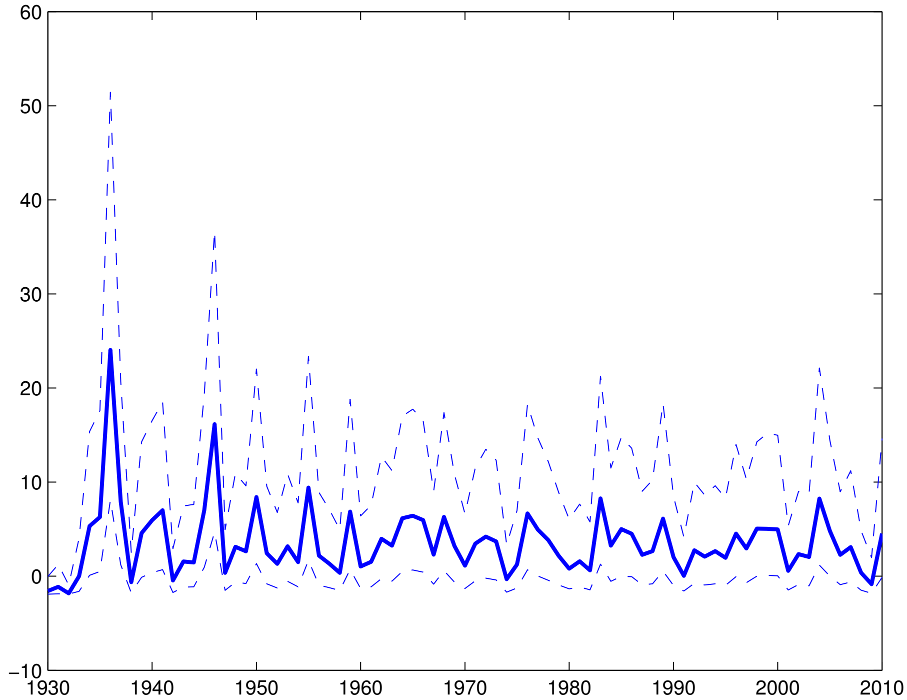

Figure 9: *Time-series of price-growth over a horizon of 3 years. Taking 2.5, 50 and 97.5th quintile of parameters and states* X_t *obtained from the MCMC, I plot expected price growth in three years using* $H(X_t, 3) = E_t\left[\exp\left(\int_t^3 \mu_P(X_s)ds\right)\right]$. *The expression for* $\mu_P(X_s)$ *is given in the appendix.*

.

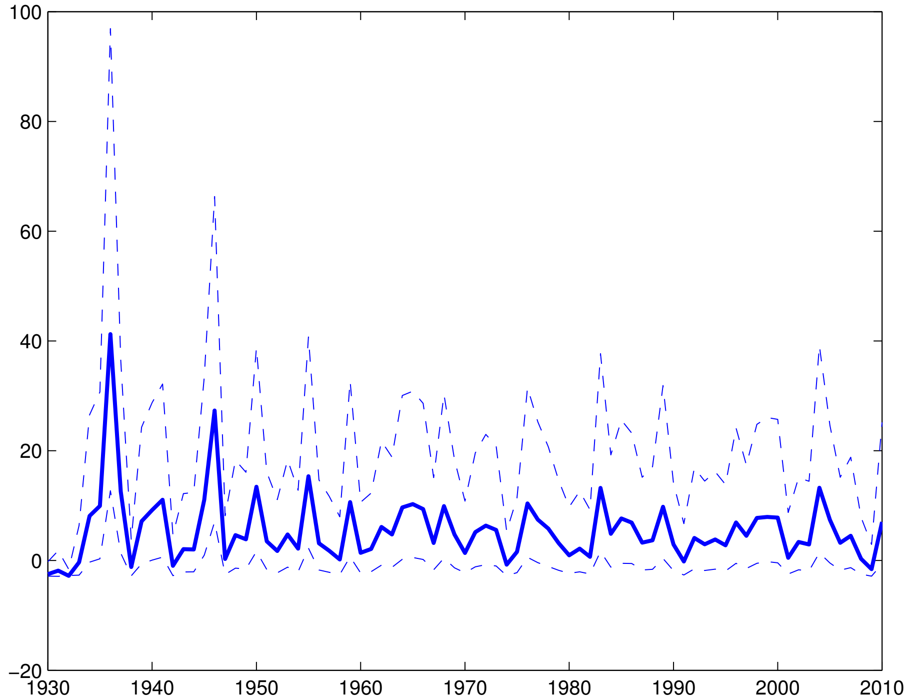

Figure 10: *Time-series of price-growth over a horizon of 5 years. Taking 2.5, 50 and 97.5th quintile of parameters and states X_t obtained from the MCMC, I plot expected price growth in five years using $H(X_t, 5) = E_t\left[\exp\left(\int_t^5 \mu_P(X_s)ds\right)\right]$. The expression for $\mu_P(X_s)$ is given in the appendix.*

.

Table IA: The following represents the parameter estimates from estimating the state-space given in (19)-(20) via a Gibbs sampler. The posterior distribution is presented in the form of 2.5-th to the 97.5-th quantiles of the simulated posterior draws from the Gibbs sampler. The data used for this sample is annual aggregate dividend from CRSP value-weighted index and BLS consumption(non-durables and services) growth in the US from 1929-2010. All nominal quantities are converted to real using CPI.

	0.025	0.25	0.5	0.75	0.975
σ_D	0.0864	0.0952	0.1004	0.1060	0.1182
σ_C	0.0174	0.0191	0.0201	0.0212	0.0235
σ_x	0.0234	0.0258	0.0271	0.0286	0.0319
κ	0.0989	0.1026	0.1044	0.1063	0.1101
λ	0.3033	0.3304	0.3440	0.3572	0.3833
μ_C	0.0108	0.0164	0.0195	0.0228	0.0293
μ_D	-0.0004	0.0139	0.0212	0.0284	0.0427

Table IB: The following represents the posterior distribution of key moments of real consumption and dividend growth implied by the model (19)-(20). The posterior distribution is presented in the form of 2.5-th to the 97.5-th quantiles of the moments computed from the parameter estimates in Table IA. The sample statistics are computed from annual aggregate dividend from CRSP value-weighted index and BLS consumption(non-durables and services) growth in the US from 1929-2010.

	Data	0.025	0.25	0.5	0.75	0.975
Mean of dividend growth	0.0231	-0.0004	0.0139	0.0212	0.0284	0.0427
Mean of consumption growth	0.0199	0.0108	0.0164	0.0195	0.0228	0.0293
Vol. of dividend growth	0.1455	0.1051	0.1133	0.1179	0.1229	0.1338
Vol of consumption growth	0.0295	0.0259	0.0280	0.0292	0.0305	0.0331
AC(1) of dividend growth	0.2127	0.1719	0.2152	0.2402	0.2684	0.3263
AC(1) of consumption growth	0.4519	0.3580	0.4307	0.4690	0.5069	0.5764
Corr of dividend and consumption growth	0.5819	0.2989	0.3437	0.3729	0.4053	0.4690

Table II: This Table documents the posterior distribution of market prices of risk given in (6). Given the full parameter distribution summarized in Table I, I compute the posterior distribution of transient volatility risk - $\gamma\sigma_C$ and long-run risk (**LR risk**) - $\frac{(\gamma-1)\lambda\sigma_x}{\kappa+\beta}$. The 2.5 to 97.5-th quantile of the posterior distribution of the two risks is presented below. Furthermore, I use $\beta = 0.001$ and $\gamma = 7.5$.

	0.025	0.25	0.5	0.75	0.975
Transient risk	0.1305	0.1432	0.1508	0.1590	0.1763
LR risk	0.4718	0.5390	0.5774	0.6186	0.7110

Table III: Below I present quantiles from the posterior distribution of endogenous quantities with $\beta = 0.001$ and $\gamma = 7.5$. Then using the full parameter distributions obtained through the Gibbs sampler, I compute the posterior distribution of instantanoues expected excess return, instantaneous volatility of cumulative return, the Sharpe-ratio, the dividend-price ratio, the risk-free rate and the volatility of the risk-free rate. For all of the following quantities, I integrate out the initial state X_t by using its stationary distribution $X_t \sim N\left(0, \frac{\sigma_x}{\sqrt{2\kappa}}\right)$. The corresponding sample statistics are obtained from CRSP Value-Weighted Market Index and the 90-day T-Bill Rate also obtained from CRSP. All nominal quantities are deflated by the CPI. Empirical estimates are obtained with GMM, and standard errors are Newey-West corrected with five lags. The data interval is annual from 1929-2010.

	Data	0.025	0.25	0.5	0.75	0.975
μ^R	0.0702(0.0177)	0.0478	0.0564	0.0614	0.0670	0.0791
σ^R	0.2011(0.0183)	0.1321	0.1417	0.1469	0.1524	0.1638
Sharpe ratio	0.3512(0.0211)	0.3335	0.3872	0.4171	0.4513	0.5247
$\frac{D}{P}$	0.0392(0.0035)	0.0316	0.0456	0.0534	0.0606	0.0760
Vol $\left(\frac{dG}{G}\right)$	0.1536(0.0211)	0.0914	0.1009	0.1064	0.1120	0.1237
r^f	0.0104(0.0078)	0.0083	0.0143	0.0175	0.0207	0.0267
$\sigma(r^f)$	0.0403(0.0059)	0.0168	0.0192	0.0205	0.0219	0.0251

Table IV: This table shows the pseudo-R^2's of the predictability relationship using (16) and (17). I restrict myself to the 0.25, 0.5 and 0.75-th quantiles of parameters given in Table I. Furthermore, I use $\beta = 0.001$ and $\gamma = 7.5$ and simulate using monthly increment by setting $dt = 1/12$.

The unconditional variance of $\bar{R}_T$ in (16) is computed by using the total variance formula - $Var(\bar{R}_T) = Var_X(E(\bar{R}_T|X_t)) + E_X(Var(\bar{R}_T|X_t))$. Starting at many different $X_t's$ drawn from its unconditional distribution, I simulate out $\bar{R}_T$ and form the inner conditional expectation and variance for each starting point. Then I perform the outer expectation and variance to compute the unconditional mean. In all, 250,000 paths are used to compute each $Var(\bar{R}_T)$. I repeat the same exercise to compute the variance of the conditional mean - $[E_t[\bar{R}_T]]$ in (17). To form pseudo-R^2, I simply divide the variance of $E_t[\bar{R}_T]$ by the variance of $\bar{R}_T$.

$z(years)$	$Quantiles$	0.25	0.5	0.75
1	$Var(\bar{R}_T)$	0.0192	0.0208	0.0224
	$Var[E_t[\bar{R}_T]]$	0.0018	0.0021	0.0026
	$pseudo-R^2$	0.0926	0.1022	0.1139
3	$Var(\bar{R}_T)$	0.0671	0.0859	0.0894
	$Var[E_t[\bar{R}_T]]$	0.0212	0.0285	0.0378
	$pseudo-R^2$	0.3164	0.3317	0.4234
5	$Var(\bar{R}_T)$	0.1568	0.1867	0.2283
	$Var[E_t[\bar{R}_T]]$	0.0731	0.1171	0.2062
	$pseudo-R^2$	0.4663	0.6272	0.9031

Table V: This table shows results of predictability regression $R_{t+z} = a + b\frac{D_t}{P_t} + \epsilon$ where $\epsilon \sim N(0,1)$ and $z = 1, 3, 5$ years. The data for this regression is generated the following way. First, I discretize the state equation (3) and simulate the growth rates X_t in monthly frequency for 720 months (roughly the size of post-war sample). Then, I compute monthly dividend growth by discretizing equation (1) and price-dividend ratio from equation (8). Using the relationship $R_{t+1} = \frac{D_{t+1}}{D_t}\frac{\frac{P_{t+1}}{D_{t+1}}+1}{\frac{P_t}{D_t}}$, I create monthly returns. From monthly returns, I compound to create 1-5 year returns and run the above predictability regressions. This is repeated 10,000 times. The parameters for the simulation are the median parameters taken from the estimation in Table I. Below, I present the median and 2.5-97.5 quantiles of the point estimate of the coefficient on dividend yield b, T-statistics of b and and R^2 from 10,000 predictability regressions from 60-year simulated data.

z(years)		median	0.025 quantile	0.975 quantile
1				
	b	0.6279	0.3192	0.9738
	T-stat	3.6120	1.4713	6.2442
	R^2	0.1863	0.0366	0.4062
3				
	b	6.6225	2.7519	10.1020
	T-stat	5.2306	1.7531	10.0113
	R^2	0.3322	0.0529	0.6457
5				
	b	38.7433	11.1834	62.1032
	T-stat	5.3320	1.3194	11.1188
	R^2	0.3491	0.0322	0.6999