Equilibrium Predictability
And Other Return Characteristics

Satadru Hore*

November 21, 2008

* Satadru Hore is at the University of Iowa, email:satadru-hore@uiowa.edu. All errors are mine.
Equilibrium Predictability
And Other Return Characteristics

Abstract

I derive equilibrium predictability results in a Bansal and Yaron (2004) one channel economy under Duffie-Epstein preferences. The dynamic properties of my model are delivered by a single persistent growth rate that serves as a common trend in aggregate consumption and dividend growth and it induces long-run risk for agents with DE preferences. The key results are delivered by an endogenously determined price-dividend ratio that is no longer exponentially affine, and as such can produce important implications about return predictability like the stochastic nature of return predicting coefficient. The setting is tractable enough to produce an analytical expression for long horizon predictability. Consistent with new findings in the empirical predictability literature, I conclude that the 1-5 year regression coefficients offer no evidence of long horizon predictability. I differ from the current equilibrium literature that shows significant predictability coefficient because I address the conditional nature of the predictability coefficient that other models imply but cannot fully assess due to the non-tractability of their solutions.
1 Introduction

This paper finds strong implications for aggregate return predictability in the one-channel Bansal and Yaron (2004) economy. 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 premia is time-varying under the assumption that the representative agent is endowed with Duffie-Epstein (DE) preferences with Elasticity of Intertemporal Substitution (EIS) equal to 1. This finding is significant because Bansal and Yaron (2004) does not have time-varying equity premia in their one-channel economy. From time-varying equity premia, I derive interesting implications for return predictability that ascribes time-series properties to the coefficient of return predictability. I find that the dividend yield and the coefficient on predictability are inversely related - a decrease in dividend yield corresponds to a faster increase in the return predicting coefficient leading to increase in equity premia. Furthermore, the setting is tractable enough to solve for long horizon regression predictability coefficient in semi-closed form. When the conditional form of the long horizon predictability coefficient is converted to the unconditional one by integrating out the underlying stochastic element, statistically speaking - all long-horizon predictability vanishes. In other words, the amount of volatility of the underlying growth rates necessary to match aggregate return volatility and premia, makes long-horizon coefficients statistically insignificant.

The theory developed here makes two potentially interesting points. First, it finds time-series implications in equity premia 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 price growth is a constant and since the one channel
economy has constant volatility in the growth rates and, by extension, in the pricing kernel, equity premia is constant and there is no predictability. In their work, the dual channel economy with stochastic volatility of the growth rate of consumption and dividend creates time-variation in the shocks of the pricing kernel which gives rise to time-varying risk-premia and predictability. In this paper, I start with DE preferences with EIS = 1, which allows me to solve for the pricing kernel by first solving the Hamilton-Jacobi-Bellman equation of the representative agent in closed form. This gives me a closed form of the pricing kernel which allows me to get an analytical expression for the PD ratio that is no longer exponentially affine. The extreme non-linearity in the log PD ratio creates time-varying volatility in returns, which gives time-varying risk-premia and predictability. Thus, I do not need stochastic volatility in consumption or dividend growth, but simply time-variation in the growth rate to get all the time-series effects in return dynamics. The economic effect of the non-linearity is significant, especially with regards to long-horizon predictability as this paper shows.

Secondly, given the closed form solution of PD ratio, I can directly solve for the long-horizon return predictability coefficient in a semi-closed form and reach different conclusions than the current equilibrium literature that addresses long-horizon predictability, like Campbell and Cochrane (1999) and Bansal and Yaron (2004). Both works imply a time-varying coefficient of predictability. Since their models are less tractable than the one here, they run simulation exercises and report OLS coefficients that are significant in the long horizon which the new empirical literature, discussed below, firmly rejects. I analyze the model implied long-horizon coefficient directly and I also find high and increasing coefficients. But, consistent with the empirical literature, I find the coefficients to be highly insignificant when I integrate out the conditioning variable using its model implied uncon-
ditional distribution.

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 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 fictitious $R^2$’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. Due to the overlapping nature of the problem, they consider multivariate regressions in the spirit of Seemingly Unrelated Regressions (SURs). This paper follows on the same spirit modelling underlying autocorrelated shocks in the dividends that show up in both dividend yields and returns, thereby creat-
ing the environment such that underlying shocks can build through time to produce high predictability coefficients in the form of expected PD ratio growth and dividend growth.

The equilibrium literature, on the other hand, show a very different outcome. In the habit persistence model of Campbell and Cochrane (1999) or Bansal and Yaron (2004), they consistently show high and positive return predicting coefficients that are statistically significant with high $R^2$'s. In the absence of an analytical solution for the long-horizon coefficient, they simulate their structural models and run univariate regressions to arrive at these conclusions on predictability. In my case, I too have a dynamic equilibrium model where the underlying growth rate of consumption and dividends driven by a single latent factor generates all the time-series effects in return, under log-recursive utility. However, my setting is tractable enough that I can have an analytical representation of the long-horizon return predicting coefficient. Due to the underlying stochastic nature of the economy, this long-horizon coefficient is also stochastic and can be decomposed into growth of PD ratio and dividend growth. Taking the parameter estimates that match aggregate return characteristics as well as macroeconomic time-series, I compute the unconditional long-horizon coefficient directly without resorting to simulation of the return series and conclude that the volatility in the PD ratio growth is too high for there to be strong, positive coefficients. Furthermore, I find that because PD ratio growth is highly volatile as horizons increase, the unconditional variance of the conditional expected return given by the long-horizon return expression has high volatility thereby generating artificially high $R^2$'s.

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 persistence. 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 fluctuation 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.

First, I estimate my dynamic model using a Bayesian Markov-Chain Monte Carlo method to get posterior estimates of model parameters and growth rates. The growth rates are filtered and smoothed estimates from the data using a Bayesian version of kalman-filter called Forward-Filtering Backward Sampling due to Carter and Cohn (1996). Then, using my full posterior estimates I replicate traditional asset pricing quantities and address predictability. A big assumption in my model is that EIS=1. This is a topic of immense debate in the macroeconomic literature. Earlier works on recursive preferences like Kandel and Stambaugh (1991) focus on EIS close to zero. Hansen and Singleton (1982), Attanasio and Weber (1989), Vissing-Jorgensen (2002) and Guvenen (2001) argue that EIS is greater than 1. To get the right asset-pricing effects in Bansal and Yaron (2004), clearly one needs EIS to be greater than 1. Campbell (1989) and Hall (1988) estimate EIS to be less than 1. To Yogo (2004) estimates the EIS parameters in several countries and finds that one cannot reject EIS=1 in every country except Canada and Switzerland. Thus, without tying my model to EIS greater than or less than one, I take the limiting case of EIS=1.

The paper is subdivided into the following parts: section 2 discusses the details of the model and establishes the predictability results. Section 3 discusses the estimation methodology. Section 4 covers the empirical findings on asset pricing quantities and equilibrium predictability.
2 The Model

2.1 Preferences and Dynamics

Regular 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.

In continuous time, the recursive utility function takes the form of stochastic differential utility. The stochastic differential utility $U : \mathcal{L}^2 \rightarrow \mathcal{R}$ is a mapping from a square integrable space to the real line and is defined by two primitive functions: $(f, A)$ where $f : \mathcal{R}^+ \times \mathcal{R} \rightarrow \mathcal{R}$ and $A : \mathcal{R} \rightarrow \mathcal{R}$. For any consumption process $C \in \mathcal{L}^2$, the utility process $J$ is the unique SDE

$$dJ_t = \left[-f(C_t, J_t) - \frac{1}{2} A(J_t) \sigma_J \sigma'_J \right] dt + \sigma_J dW_t$$

with boundary condition $J_T = 0$. The different components are - $J_t$, a continuation utility for the agent given consumption $C_t$, $f(C_t, J_t)$ is an ordinal map of date $t$’s consumption and continuation utility, $A(J_t)$ is a measure of local risk-aversion and $\sigma_J$ is the volatility of the utility process. If given an initial consumption $C_t$ and as long as the solution of
the above SDE is well-defined, the utility at time $t$ is defined as $U(C_t) = J_t$. Under certain technical conditions, the above SDE is well-defined and hence the utility function exists. The function $U$ is monotonic and risk-averse for $A \leq 0$. Given an $f$ and two functions $A^*$ and $A$, let $U^*$ and $U$ be the two utilities corresponding to the aggregators $(f, A^*)$ and $(f, A)$. If $A^* \leq A$, then $U^*$ is more risk-averse than $U$, i.e. any consumption stream rejected over a deterministic consumption path by $U$ will also be rejected by $U^*$. A convenient normalization that produces an ordinally equivalent utility function is achieved by setting $A = 0$, which means the above SDE solves $E_t[dJ_t] + f(C, J) = 0$ for normalized aggregator $(f, 0)$. The normalization is useful because it produces a much simpler Bellman equation to be solved than if $A \neq 0$. Fortunately, there exists a transformation from $(\bar{f}, A)$ to $(f, 0)$ such that the utilities generated from both will be ordinally equivalent. Further discussion of the aggregators and the normalization that leads to an ordinally equivalent representation of the aggregators is given in Duffie-Epstein (1992).

The utility function that is considered here is due to Duffie and Epstein which is a continuous time counterpart of Kreps-Porteus and Epstein-Zin-Weil. The normalized aggregator corresponding to $A = 0$ is

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

where $\psi$ is the EIS, $\beta$ is the discount rate and $\gamma$ is the risk-aversion. Assume the representative investor is endowed with a log-recursive utility, which is the special case of above preferences with $\psi = 1$. Thus the normalized aggregator $f$ for $\psi = 1$ is the limit of the
above utility function as $\psi \to 1$.

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

Assume that consumption and dividend growth jointly follow a geometric path with time-varying growth rate $X_t$, which represents a small predictable component in dividend growth.

$$\frac{dD}{D} = (\mu_D + X_t)dt + \sigma_D dW_D$$  \hspace{1cm} (1)

$$\frac{dC}{C} = (\mu_C + \lambda X_t)dt + \sigma_C dW_C$$  \hspace{1cm} (2)

$$dX_t = -\kappa X_t dt + \sigma_x dW_X$$  \hspace{1cm} (3)

where the brownian motion shocks are all uncorrelated. This formulation has its origin in Abel (1999) which posits that consumption is a claim to a fraction of aggregate dividend and the rest is distributed as debt-payments in the economy. In this dynamic setting, $X_t$ represents a mean zero latent growth rate shock that captures the time-series variation in the expected growth rate of aggregate dividends, and $\lambda < 1$ has its effect of slowing down corresponding expected consumption growth rate. In the language of cointegration, the expected growth rate $X_t$ represents the cointegrating residual and should be stationary. A high (low) growth rate shock $X_t$ on aggregate dividends produces a milder expected growth rate shock on aggregate consumption due to $\lambda < 1$ which is confirmed by the time-series dynamics of real consumption and dividend growth in the post-war years shown in Figure 1. Real dividend growth has a lot of variation across time, whereas the corresponding real consumption growth is rather smooth suggesting that $\lambda < 1$. Note that unless $\lambda = 1$, the above specification rules out cointegration between dividends and consumption.
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-premia. 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

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

where $\mathcal{D}J$ 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 infinitely long-lived and has reached the equilibrium over time.

**Proposition 2.1.1** 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

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

where

$$u_1 = \frac{\lambda(1-\gamma)}{\kappa + \beta}$$
$$u_2 = \frac{1-\gamma}{\beta} \left[ \mu_C - \frac{1}{2} \gamma \sigma_C^2 + \frac{\lambda^2 (1-\gamma) \sigma^2}{2(\kappa + \beta)^2} \right]$$
2.2 Asset Pricing

For the purpose of asset pricing, I need to derive the pricing kernel which corresponds to the special case DE preferences with EIS=1. Duffie and Epstein (1992) show that the pricing kernel for stochastic differential utility, \( \Lambda_t \) is given by \( \Lambda_t = \exp(\int_0^t f_s ds) f_c \).

**Proposition 2.2.1** The pricing kernel is given by

\[
\frac{d\Lambda}{\Lambda} = -r^f_t dt - \gamma \sigma_C dW_C - \frac{(\gamma - 1)\lambda}{\kappa + \beta} \sigma_x dW_X
\]  

(6)

where

\[
r^f_t = \mu_C + \lambda X_t + \beta - \gamma \sigma_C^2
\]  

(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, \( \gamma \), increases, whereas in power utility I would need \( \gamma \) 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 \( \gamma \)-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 risk-free rate owing to the fact that I have set EIS=1. The proposition that risk-free rate decreases in risk-aversion uniformly in the log-recursive case is not surprising. Recall, that in the log-recursive case \( \gamma > 1 \) is sufficient to generate preference for early resolution of uncertainty. If risk-aversion is increasing, the agent is willing to settle for a lower certainty equivalence in the future which reduces the risk-free rate.
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:

$$\frac{(\gamma-1)\lambda}{\kappa+\beta} \sigma_x = \frac{\lambda}{\beta} \sigma_x$$

which measures a change in the value function of the agent with respect to the growth rate $X_t$. In these recursive preferences, the value function is embedded within the aggregator, 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 $\gamma$, but the effect is magnified due to $\kappa$ and $\beta$ in the denominator. Recall, that the stationary distribution of $X_t \sim N \left(0, \frac{\sigma^2}{2\kappa} \right)$. Thus, as $\kappa$ decreases, in other words as the growth rate becomes more persistant, the volatility of these growth rates increase and an agent exposed to long-run risk from these persistant growth rates seek higher compensation for this risk. Notice, that the absolute magnitude of the size of long-run risk can be astronomical. Given a modest risk-aversion parameter, say $\gamma = 7.5$, a small $\kappa$, say $\kappa = .02$ as in Bansal and Yaron (2004), can greatly magnify the price of this risk for long lived agents with already small $\beta$. In comparison, the traditional transient component of risk is tiny for $\gamma = 7.5$. In fact, Table 2 shows that post-war US data on real dividends and consumption imply a market price of long-run risk to be roughly 7-9 times the size of the traditional transient consumption risk under reasonable preference and aggregate parameters.

Given the pricing kernel of the stochastic differential utility, I establish the equilibrium price-dividend ratio and return dynamics.
Proposition 2.2.2 Equilibrium price-dividend ratio is given by

\[ \frac{P_t}{D_t} = G(X_t) \quad (8) \]

where \( G(X_t) = \int_1^\infty \exp(P_1(\tau)X_t + P_2(\tau))d\tau \), 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

\[ dR = \frac{dP + Ddt}{P} - r_f dt = \mu^R t dt + \sigma_D dW_D + \frac{G_X}{G} \sigma_x dW_x \]

where equilibrium expected excess return is

\[ \mu^R = \frac{\lambda(\gamma - 1)}{\kappa + \beta} \frac{G_X}{G} \sigma_x^2 \quad (9) \]

and the volatility of cumulative return given by

\[ \sigma^R_t = \sqrt{\sigma_D^2 + \left( \frac{G_X}{G} \sigma_x \right)^2} \quad (10) \]

This is the central result in the paper. Unlike the one-channel economy of Bansal and Yaron (2004), the PD ratio in this one-channel economy is no longer exponentially affine but is highly non-linear in the expected growth rate. The non-linearity of the growth rate in the log PD ratio is responsible for generating all time-series results in this paper. 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 \( G = \frac{P}{D_t} = \exp(a + bX_t) \), where \( a \) and \( b \) are constants, then \( \text{Vol} \left( \frac{dG}{G} \right) = b \). Thus, if market price of risk from \( X_t \) is also a constant, then risk premia will be a constant thus eliminating any time-series phenomenon.
in expected returns or any dynamic predictability relationships. In the Bansal and Yaron (2004) economy, there is time-series effect in returns 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 in the one-channel case (just like Bansal and Yaron (2004)), but the non-linearities in the log PD ratio creates time-varying volatilities in prices which creates time-varying equity premia. Since the PD ratio is not exponentially affine, I can generate rich time-series dynamics in aggregate market returns with interesting predictability implications described below. The necessary transversality condition for the integral to converge is given in the proof of the proposition which ensures that the PD ratio is finite. Notice that if consumption and dividend were cointegrated, i.e. if \( \lambda = 1 \), the PD ratio is constant because then \( P_1(\tau) = 0 \). Thus, I need consumption and dividends to be un-cointegrated for the time-series effects to go through. The price to consumption ratio (total wealth), however, is completely uninteresting. The risk-free rate cancels out the time-varying component of expected consumption growth since \( \text{EIS} = 1 \), and the price-consumption is thereby a constant. Hence, returns to aggregate wealth is also a constant - \( 2C \).

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 the impact on price due to stochastic growth rates which is responsible for long-run risk. The equilibrium asset pricing literature has struggled with the excess volatility puzzle because canonical models imply that volatility of return is the volatility of dividend growth, where the former is much larger than the latter in the data. Here, I reconcile the excess volatility puzzle with the second term that is due to long-run risk. Notice that if \( \lambda = 1 \), then this second term from long-run risk would disappear because price-dividend ratio, \( G(X_t) \), would be a constant.
like in canonical models. 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) on the stock seeks compensation for only long-run risk since the correlation between all the Brownian motion terms are shut off. It is straightforward to incorporate those kind of risks from correlation, but for brevity I focus only on the long-run risk generated from non-linearity in the price-dividend ratio. Notice that

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

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 positive. Clearly, if $\lambda = 1$, then the effect of long-run risk vanishes since the PD ratio is constant for $\lambda = 1$ and equity premia vanishes because the long-run risk is the only component the agent is compensated for in this economy. But, as long as $\lambda < 1$, expected excess return is always positive and time-varying because of non-linearity in the log price-dividend ratio. The rich time-series implications of this model is now investigated.

## 2.3 Time-Series and Predictability

The non-linearity in the log price-dividend ratio in the one-channel Bansal and Yaron (2004) economy 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

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

expected return is increasing in $X_t$.\(^1\) Therefore, we have a situation where PD ratio, $G(X_t)$, rises with growth rate as well as expected returns. In response to a positive shock from underlying economic growth rates, expected dividend growth increases which increases prices relative to dividends as long as $\lambda < 1$. But, positive economic growth rates also increase expected returns. 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 at a faster rate than 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

$$
\mu_t^R = \left[ \frac{(\gamma - 1)\lambda}{\kappa + \beta} \sigma^2 G_X \right] \frac{D_t}{P_t}
$$

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 of the coefficient on dividend yield, $G_X$, has the property that $(\frac{G_X}{G})_X > 0$. This ensures that as $G(X_t)$ increases (dividend yield decreases) as $X_t$ increases, $G_X$ increases faster than $G(X_t)$ which “pulls up” a diminishing dividend yield to produce higher expected return. It is shown pictorially in Figure 5. This makes the coefficient of predictability itself time-varying - a fact shown in Lettau and van Nieuwerburgh (2008). This stochastic nature of

\(^1\)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.
the predictability coefficient is sorely missing in the predictability literature. It helps us to understand how when the economy is good we can see both high prices and high returns. It helps make returns stationary - when dividend yield changes, the return predicting coefficient moves in the opposite direction as dividend yield. 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. The inverse relationship between dividend growth and dividend yield is precisely the “non-barking dog” of Cochrane (2008), i.e. the failure in the data to confirm this negative relationship is a lot more significant than the absence of return predictability. Now I move on to predictability in the long horizon. I derive an analytical expression of the conditional mean relationship between dividend-price ratio at time $t$ of conditional expected return at time $T$.

2.3.1 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

$$R_T = \frac{P_T + \int_t^T D_r \, dr}{P_t}$$

(12)

Note that this is different from the instantaneous excess return dynamics developed in (2.2.2), where I use $dR_t = \frac{dP_t + D_t \, dt}{P_t} - r^f_t \, dt = -E_t \left( \frac{dP}{P}, \frac{dA}{A} \right) + \cdots dW$. The latter expression, integrated forward to produce $R_T$, is wholly unsuitable to analyze long-horizon cumulative
returns. First of all, the instantaneous cumulative return dynamics is of a $dt$-period return from $t$ to $t + dt$ with dividends $D_t$ and risk-free rate $r^f_t$ held constant at time $t$. The expected growth rate $X_t$ also stays constant, and I only account for price change due to $X_t$. Then dividends are ploughed back in at a constant rate $D_t dt$ and the risk-free rate is also subtracted at a constant rate $r^f_t dt$ to produce excess return. Integrating forward this quantity will not address the fact that there are dynamic relationship 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 (12) where I accumulate dividends from $t$ to $T$ and also consider the intermediate shocks from dividends 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

$$P_T = P_t \exp\left(\int_t^T \left[\mu_P(X_s) - \frac{1}{2}\sigma_P(X_s)\sigma_P(X_s)'\right] ds + \int_t^T \sigma_P(X_s) \cdot dW\right)$$  \hspace{1cm} (13)$$

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. A big distinction between this price growth and the price growth from the instantaneous return relationship is that the latter, (9), will show long horizon expected returns that are always positive because we are simply adding equity-premia that is always positive, but the relationship in (13) can produce either pos-
itive or negative expected return as we will see below. Similarly, dividend growth can be written as

\[
D_r = D_t \exp \left( \int_t^T \left( X_s - \frac{1}{2} \sigma_D^2 \right) ds + \int_t^T \sigma_D dW_D \right) \tag{14}
\]

Substituting them both into (12), I can write total return from \( t \) to \( T \) as

\[
R_T = \left[ G(X_t) \exp \left( \int_t^T \left[ \mu_P(X_s) - \frac{1}{2} \sigma_P(X_s) \sigma_D(X_s) \right] ds + \int_t^T \sigma_P(X_s) \cdot dW_s \right) + \int_t^T \exp \left( \int_t^T |X_s - \sigma_B|^2 ds + \int_t^T \sigma_D dW_D \right) dr \right] \frac{D_t}{P_t} \tag{15}
\]

This expresses cumulative return over horizon \( z = T - t \) as a function of 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 PD ratio growth from \( t \) to \( T \) and the second term is the growth in dividends (fixing initial dividend \( D_t = 1 \)). 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. 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 (15), has a more tractable form without the Brownian shocks. First, observe that the conditional expectation of future dividends has a closed-form solution.

**Lemma 2.3.1** 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 (12).
Proposition 2.3.1 The price process in (13) implies \( E_t[P_T] = P_t H(X_t, z) \) 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 (2.3.1)

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

(16)

Notice, that the expression \( H(X_t, z) \) is conditionally known at time \( t \) and is a function only of \( X_t \) and horizon \( z \). To see it, one can write \( E_t[P_T] \) (assuming that the expectation is finite) as \( E_t[P_T] = P_t H(X_t, z) \) from a direct application of Feynman-Kac. \( H(X_t, z) \) satisfies a second order partial differential equation that depends on \( X_t \) and \( z \). It has an unique solution given a set of boundary conditions. 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-linearities 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 \( E_t\left[ \int_t^T \mu_P(X_s) ds \right] \) for every initial point \( X_t \).

The \( z \)-horizon return predictability coefficient \( \alpha(X_t, z) \) is composed of two parts. There is a dividend growth component and then a PD ratio growth component as a dynamic response to dividend growth rates and dividend shocks. The dividend growth can be solved in closed-form whereas price change is dictated by the functional relationship \( H(X_t, z) \) which can be solved numerically by the integral given in Proposition (2.3.1). 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 asymptotics. The structural relationship here suggests that the slope coefficient on these long (or short!) horizon regressions are themselves stochastic with time-series properties discussed before, 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 (15) 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 which he analyzes by using the Functional Central Limit Theorem. To convert the conditional expectation relationship into a percentage return form, I simply subtract one and focus on the quantity.

$$E_t[R_T] = G(X_t)(H(X_t, z) - 1) + \int_t^T \exp(A(r - t)X_t + B(r - t))dr \frac{D_t}{P_t}$$ (17)

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 long horizon. 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 whether the $R^2$-s are increasing over time. 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 strong biases arising from regressions on lagged dependent variables. In fact, they find 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 (16), provides a powerful estimate of long-horizon return predictability. To gauge the power of the relationship (16) in the form of $R^2$, I ask the question - How much of the unconditional variance of $R_T$ can be explained by the unconditional variance of the conditional mean relationship in (16)? Thus, to infer the power of the long horizon estimate (16), the $R^2$ in my setting is simply $R^2 = \frac{\text{Var} \left( \alpha(X_t, z) \right)}{\text{Var}(R_T)}$, where $\text{Var}$ denotes unconditional variance. Written in this way, it is clear that unconditional variance of the conditional mean relationship in (16) does not have any easy form. Since the coefficient of the conditional mean $\alpha(X_t, z)$ is itself a function of $X_t$ that also impacts the dividend yield $\frac{1}{G(X_t)}$, I cannot simply pull it out of the expression that I can if I treated it as a constant. Empirical works that treat the return predicting coefficient as a constant misses this extra uncertainty behind the coefficient.

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 PD ratio $G(X_t)$ or the current expected growth rate $X_t$, but because of the impact of persistant growth rates producing persistant PD ratios across time. Naturally this persistance will be stronger in the short-run between $T_1$ and $T_2$ than say 4-year regressions $T_1$ and $T_4$, just because growth rate paths and the corresponding price paths will be more correlated in the short horizon than in the long horizon. Empirically, BRW (2008) has found that this correlation between the return predicting coefficients is quite significant.

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 closed-form estimate of the conditional mean of long-horizon regression with explicit expression for the long-horizon predictability coefficient. The result shows a strong time-series dependence of the return predicting regression coefficient and dividend yield rendering inference drawn from pure OLS based exercises rather biased. In addition, I can simulate $R^2$s in this setting to test the power of the equilibrium predictability relationship given in (16).

3 Empirical Methodology

3.1 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 $p(\theta, X_{t,T})$ conditional on the data on consumption and dividend growth. We will follow a Markov Chain Monte Carlo (MCMC) algorithm that will draw them conditionally on each other

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

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

$$
\begin{bmatrix}
g_{t+1}^d \\
g_{t+1}^c \\
X_{t+1}
\end{bmatrix} =
\begin{bmatrix}
(\mu_D + X_t) \\
(\mu_C + \lambda X_t)
\end{bmatrix} +
\begin{bmatrix}
\sigma_D & 0 \\
0 & \sigma_C
\end{bmatrix}
Z_1
$$

(18)

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

(19)

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 $X_{1,T}$ conditional on the rest of the parameter space, $\theta$, and the full time-series of dividend ($g_{1,T}^d$) and consumption ($g_{1,T}^c$) 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 $\theta$, and I draw the full time-series of $X_{1,T}$ given the full time-series of dividend and consumption growth.

Then, my goal is to draw the parameter set $\theta|X_{1,T}, g_{1,T}^d, g_{1,T}^c$ conditional on the full time-series of the growth rates $X_{1,T}$ which I have estimated in the above step using FFBS, and the data on consumption and dividend growth. 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 - $\theta_i|\theta_{-i}, X, y$, where $\theta_{-i}$ 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 parameters $\{\mu_C, \lambda\}$ are obtained from running a Bayesian regression of consumption growth rate $g_{t+1}^C$ on the filtered and smoothed growth rates $X_t$, and drawn from a multivariate normal. The parameter $1 - \kappa$ is obtained by running an AR(1) regression on the filtered and smoothed growth rates $X_t$, and drawn from an univariate normal. $\sigma_C$
and $\sigma_x$ are drawn from a chi-squared distribution conditional on knowing the parameters $\{\mu_C, \lambda\}$ and $1 - \kappa$, respectively. Finally, from the dividend growth rate $g_{1,T}$, I draw $\mu_D$ and $\sigma_D$ from an univariate normal and chi-squared distribution, respectively. The exact parameters of these posterior distributions is discussed in detail in Allenby, McCulloch and Rossi (2005).

3.1.1 Priors

The strength of the Bayesian mechanism is the ability to specify prior information on the growth rate $X_t$ since it is not directly observable. Prior belief on the parameters of $X_t - \kappa$ and $\sigma_x$, 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 and smooth. To obtain persistant and smooth states, I impose a prior on $1 - \kappa \sim N(0.95, 0.1^2)$ and a prior on $\sigma_x$ such that the “noise-to-signal” ratio between dividend growth and expected dividend growth is 10. This prior specification greatly smoothes the latent state relative to the actual growth rates as will be showed below. The prior sensitivity is checked with other specifications and available upon request. Finally, since the theory heavily relies on $\lambda < 1$, I propose the prior $\lambda \sim N(0.5, 0.5^2)$ which is not too informative since it implies prior subjective belief of $\lambda \in \{-0.5, 1.5\}$.

3.2 Data

I restrict myself to postwar US data from 1948-2006 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 repur-
chases. 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.

4 Empirical Findings

The Gibbs sampler produces simulations of parameter values from their posterior distributions. The estimates from the state-space estimation (18)-(19) is reported in Table 1 in five different quintiles from 2.5-97.5-th quintile. The posterior densities are reported in Figure 2. The effect of the prior information on $\kappa$ and $\sigma_x$ (through the “noise-to-signal” ratio) is significant, as the posterior distribution in both cases is very close to the prior which helps to generate persistent and smooth growth rates that match the time-series behavior of the consumption and dividend growth quite well as is shown in Figure 3. It shows the 2.5-97.5-th quintile of the time-series of expected consumption and dividend growth rates against the actual growth rate of these quantities. It is clear that imposing the prior restriction on the latent growth rate has not had any deleterious effect in terms of matching the time-series of the underlying macroeconomic series. The most important parameter that governs the time-series effects of the asset pricing quantities is $\lambda$. If $\lambda = 1$, then dividend and consumption growth will be cointegrated and there will be no time-series effect. Notice, that the prior restriction on $\lambda$ is fairly flexible as it spans the interval $(-0.5, 1.5)$ that includes 1. The 2.5-97.5-th quintile of the posterior distribution of $\lambda$ is between (0.16, 0.28), which suggests that the data has had a substantial effect to overwhelm the prior restriction. Incidentally, the Bansal and Yaron (2004) calibration of the “$\phi$”-parameter in their one-channel
4.1 Market Prices of Risk

To focus on asset pricing, I pick the following preference parameters - time discount parameter $\beta = .01$ and risk-aversion $\gamma = 7.5$ following Mehra and Prescott’s (1985) suggestion that we should be looking for risk-aversion less than 10 to match asset pricing facts. The posterior estimates of the market prices of risk is in Table 2. 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 an astronomical margin. Whereas the posterior mean of the price of transient risk is 0.08, the posterior mean of the price of long-run risk is 0.43. 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. The reason for such high price of long-run risk is clear from equation (6). Small $\kappa$ and $\beta$, can greatly magnify a modest risk-aversion parameter in the long-run risk term. Thus, the persistant expected growth rate of consumption and dividend growth makes a strong contribution to the high market price of long-run risk.

4.2 Asset Pricing

This subsection shows that the parameter estimates estimated above from the simple state-space setting (18)-(19) can match key unconditional asset pricing quantities, a la Bansal and Yaron (2004)’s one-channel economy. The time-series implication is addressed in the next
section. The magnitude of the same quantities computed here is much more pronounced than Bansal and Yaron (2004)’s due to the non-linearity in the price-dividend ratio discussed above. By 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 change in dividend-price ratio, risk-free rate (7) and the Sharpe Ratio. Notice, that since all of these quantities depend on the growth rate $X_t$, I integrate it out by using the stationary distribution of $X_t \sim N(0, \frac{\sigma^2_\mu}{2\xi})$ to produce unconditional estimates. Figure 4 plots the posterior distributions of these equilibrium quantities, with a line indicating the unconditional value of these endogenously determined quantities observed in the data and Table 4 reports the 2.5-97.5-th quintile of these quantities.

The posterior distribution of expected return is between 7.01-11.98%, whereas the average return on the market portfolio in the post-war sample is 7.44%. Similarly, volatility of return in my model is between 14.3-22.97%, whereas the sample standard deviation of the return on market portfolio is 16.62%. The lion’s share of the overall volatility is due to the volatility of the change in PD ratio. Notice, the volatility of return can be written as

$$\sigma^R_t = \sqrt{\sigma^2_D + \left( \frac{G X}{G} \sigma_x \right)^2}$$

which shows that the volatility of return can be decomposed into volatility of dividend growth, and the volatility of the change in price-dividend ratio due to changes in the growth rate of dividends. The posterior interval of the latter is 12.96-22.17%, whereas the empirically observed volatility of the percentage change in dividend-price ratio is 12.92% in the post-war US data.
Furthermore, the empirically observed Sharpe ratio is 44.82%, whereas the posterior distribution of the Sharpe Ratio here is between 38.78-74.52%. The maximal Sharpe ratio defined in Hansen and Jagannathan (1991) as the conditional volatility of the pricing kernel is between 0.39-0.93, which includes the Bansal and Yaron (2004) estimate of 0.73. Clearly, the parameters of the state-space that match the time-series of consumption and dividend growth do very well in matching market return characteristics. Notice that since all the correlations between the Brownian motion terms are shut off, the only source of risk that is priced in this economy is due to long-run risk. Thus, long-run risk from persistent growth rates along with DE preferences are enough to explain key quantities of aggregate equity returns.

The posterior distributions of dividend yield and the risk-free rate also contain the unconditional means of those quantities that is observed in the data. Another important endogenously determined quantity is the predictability coefficient from (11), which is discussed next.

4.3 Long Horizon Predictability

The standard iid view of the world implies that returns are unpredictable. However, standard regression tests like Shiller (1981) 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, which doesn’t fully reveal the conditional nature of
their underlying dynamical system.

In my case the long-horizon coefficient can be solved in a semi-closed form setting and is shown to be highly stochastic which has ramifications with regards to long-horizon predictability. Furthermore, I can isolate the effect of PD ratio growth and dividend growth on the coefficient. Both of these quantities are conditionally known and I can compute them without running any regressions, and can integrate out the conditioning information by using the unconditional density of the growth rate - \( X_t \sim N \left( 0, \frac{\sigma^2}{\sqrt{2\pi}} \right) \). I also compute how informative this conditional mean relationship is by computing a pseudo-\( R^2 \) ratio that measures the size of the unconditional variance of the conditional mean with respect to the unconditional variance of the long-horizon predictability relationship.

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

\[
\mu^R_t = G_X \frac{\lambda(\gamma - 1) \sigma^2 D_t}{\kappa + \beta P_t}
\]

where \( \frac{D_t}{P_t} = \frac{1}{G(X_t)} \). The bottom row of Table 4 presents the posterior distribution of the predictability coefficient, which is large and positive. The coefficient is between 1.66-4.98, which reflects the instantaneous effect of dividend yield and expected return. Taking the median value of 2.41, the number suggests that when dividend yields rise by one percentage point, prices rise - not fall, by another 1.41 percent, instantaneously. However, the return predictability coefficient itself has interesting time-series properties which is illustrated through Figure 5.

Taking the median values of parameters from Table 1, I compute expected return, dividend yields and the stochastic component of the predictability coefficient \( G_X \) across high and low growth rates \( X_t \). As \( X_t \) increases, clearly expected dividend growth increases. In response, the middle graph in Figure 5 shows that PD ratio increases, or dividend yield
decreases. However, expected return in the top graph also increases. The only way that is feasible is if the return predicting coefficient increases (goes in the opposite direction of dividend yield) and at a faster rate. The bottom graph of $G_X$ suggests precisely that. We have already established through Cauchy-Schwartz inequality that $(\frac{G_X}{G})_X > 0$, which suggests that the return predicting coefficient “pulls up” (“pushes down”) a decreasing (increasing) dividend yield to make expected return to be increasing (decreasing) in the growth rates. Missing this stochastic nature of the return predicting coefficient and its correlation with the regressor will bias the results severely even with usual standard error corrections if I were to simply use standard OLS.

Given that in my setting I can isolate the predictability coefficient, I can analyze it in detail and especially over long-horizon. Notice that Cochrane (2008) finds that high one-year return forecasting coefficient comes with low coefficient on autocorrelation of dividend yield resulting in low long horizon predictability coefficient. However, in my model high persistence in underlying growth rates result in high persistence in autocorrelation in dividend yield \(^2\) and thus I should be able to generate high long horizon predictability coefficient.

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

$\alpha(X_t, z)$ is the $z$-horizon predictability coefficient that depends on the current growth rate $X_t$ and the horizon $z = T - t$. The first component of the growth rate reflects PD ratio growth from $t \rightarrow T$, and the second reflects dividend growth. Notice, that both come about due to the same long run shock that impacts expected dividend and consumption growth. Furthermore, all analysis is simply based on computing the endogenously deter-

\(^2\)measured by $-\frac{\mu_X}{\sigma}$ in the appendix and shown through simulation.
mined expected PD ratio growth and expected dividend growth, conditional on $X_t$ (and then integrating it out). Taking the 25, 50 and 75-th quintile of parameters of the underlying process that is estimated via the Gibbs sampler, I compute dividend growth, PD ratio growth and the return predicting coefficient unconditionally.

Table 5 summarizes long horizon regression coefficients. The first row on top shows the starting PD ratio at 25, 50 and 75th percentile of the underlying parameters from Table 1. For each set of parameters and by picking $X_t$, I compute the PD ratio growth, $G(X_t)(H(X_t, z) - 1)$, and dividend growth, $\int_t^T[\exp(A(s)X_t + B(s))]ds$, across the different horizons and then add them together to produce the predictability coefficient - $\alpha(X_t, z)$. Then, I integrate out the initial state by drawing $X_t$ from its unconditional density to produce the unconditional counter-parts of the same quantities.

First, let me focus on dividend growth. Taking, the figures from the median parameter estimates, the median dividend growth in years 1-5 is $\{1.02, 1.06, 1.08, 1.12, 1.16\}$ (starting with $D_t = 1$) but with very wide standard deviations in each case making expected dividend growth completely unpredictable. Roughly, we cannot reject the fact that dividends grew at a pace of 3% annually, which is the unconditional average dividend growth in the post-war US data. However, that 3% growth has a lot of uncertainty built into it - the standard deviation is roughly 13% in the data for 1 year dividend growth. That is also reflected in the high standard deviation of expected dividend growth in the simulations here - $\{0.04, 0.08, 0.12, 0.17, 0.23\}$ in 1-5 years.

Now, let me address price growth. To give an intuition of what this term looks like, Figure 6 plots the $H(X_t, z) - 1$ function for different values of $X_t$. Notice, PD ratio growth is clearly increasing in $X_t$. It is mostly positive - high current expected growth rate produces high expected PD ratio growth. However, notice that if the current state is large
and negative, then the PD ratio growth would also be negative - in other words, bad states of the world might result in a decrease of the future PD ratio, which can only come from a price drop because expected future dividend is dropping. Taking the median estimates from Table 5 again, PD ratio growth is \( \{0.06, 0.14, 0.22, 0.32, 0.41\} \) which shows that the unconditional distribution of PD ratio growth is increasing over time, roughly at a 6-8% growth rate yearly. However, the standard deviation of these growth rates are enormous - \( \{0.06, 0.12, 0.19, 0.25, 0.34\} \) which shows what Figure 6 already predicted - that the negative expected price growth in bad times will impact the unconditional estimate of PD ratio growth when I integrate out the current state. Thus, not only is dividend growth unpredictable, but also PD ratio growth. Notice, this negative expected PD ratio growth wouldn’t have been possible if I simply integrated forward the instantaneous expected return in (2.2.2) because the equity-premia - the drift of cumulative return is always positive, and if I accumulate positive growth rates I will always get a positive PD ratio growth. Once I fully take into account the dynamic interaction between growth rates, dividends and risk-free rate, as in (13), then these interactions produce very different results as is reflected through negative price growth in Figure 6 and in Table 5. But, this makes complete intuitive sense. If the underlying growth at time \( t \) is poor and persistent, then bad times will prevail in the future, expected dividend growth is supposed to be poor in the future and price should decrease. On the other hand, adding up positive premia out into the future will always produce positive expected price growth, even though underlying growth rates are negative - which is quite contrary to logic.

The combined effect from dividend and PD ratio growth renders the 1-5 year return predicting coefficients useless. The unconditional long-horizon coefficients \( \alpha(z) \) become \( \{4.27, 8.53, 13.07, 18.08, 22.15\} \). They are clearly large and increasing. BRW show that
when these long-horizon coefficients are scaled by the 1-year coefficient, they roughly scale with time showing no extra information in long-horizon regression than in short-horizon ones. That is clearly the case here. When I divide through the $\alpha(z)$’s by the 1-year coefficient, I get $\{1.00, 2.0, 2.82, 4.00, 5.19\}$, which confirms what BRW finds. Furthermore, the standard deviation behind these coefficients are large - $\{4.86, 10.50, 23.33, 25.59, 45.24\}$. While the coefficients are increasing, the uncertainty behind them is increasing even faster, making these long horizon coefficients irrelevant. Notice in Campbell and Cochrane (1999) and Bansal and Yaron (2004), they show long-horizon regression coefficients that are increasing and highly significant. Although their models are dynamic in nature and should have these predictability coefficients as time-varying, they run regressions on simulated data making these coefficients constants! That eliminates the stochastic element in their predictability coefficients, which leads to the erroneous conclusion that coefficients are always large, positive and significant. I show here that once we take into account the stochastic element of the predictability coefficient, as shown in Valkanov (2003), the unconditional estimate of these coefficients are rather insignificant. What helped in my setting is that due to the simplicity of the underlying state-space and preferences, I can actually solve for the long-horizon predictability coefficient in an analytic expression that is easy to evaluate. The difficulty of expressing this long-horizon coefficient as a function of underlying shocks in other equilibrium models like Bansal and Yaron (2004) and Campbell and Cochrane (2008) forces them to run regressions on simulated data, thus ruining the time-varying nature of the underlying coefficients that their models imply.

The pseudo-$R^2$’s in my setting are actually increasing and economically significant! Both variances - variance of unconditional return and variance of the conditional mean given by the predictability relationship in (17) are both increasing, but the variance of the
conditional mean is increasing faster over the horizon. The median $R^2$'s range from 6% for the 1-year regression to 23% for the 5 year regression. It is not surprising that even though I do not find any strong long horizon predictability, my $R^2$'s are still increasing. That is essentially the point of Goetzmann and Jorion (1993). In their simulations, the underlying stock follows a random walk and dividends don’t, but when running regressions of returns on dividend yields they find evidence of high and misleading $R^2$'s. In my case, the $R^2$ increases mostly due to the volatility of the growth in PD ratio. PD ratio growth on average is increasing, but the uncertainty behind it increases faster making the numerator of the $R^2$ relationship rise faster than the overall unconditional variance of return in the denominator.

Thus, I address a puzzling inconsistency in predictability that shows up in equilibrium asset pricing models. Whereas more and more empirical works cast doubt on the predictability relationship, equilibrium models continue to show strong patterns of predictability. In a model that otherwise can replicate basic features of post-war US aggregate return statistics, I show that the conditional and unconditional predictability relationships are quite different. The underlying stochastic economy that connects dividend yields to expected returns has a stochastic coefficient which in the long-horizon is simply the sum of expected dividend and PD ratio growth. Conditionally, the regression coefficient is increasing and significant and the regressions have high $R^2$'s, but due to the stochastic nature of the coefficient, any unconditional result that integrates out the latent state - which was responsible for the time-series predictability result - will yield regression coefficients that are insignificant due to high volatility in price growth.

The results here pose an econometric challenge for empirical works on predictability that is hardly addressed. If the underlying economy moves according to these latent low-

36
frequency shocks then any equilibrium setting that prices these shocks will necessarily produce a predictability relationship whose coefficient is time-varying. Estimating it like an unconditional model will be deeply flawed from a structural setting. This is the reason why equilibrium models when they run predictability regressions all show positive and significant predictability coefficients with high $R^2$‘s. The unconditional aspect of the model is probably better addressed via a state-space setting, where the coefficients are time-varying and they should be filtered out from the predictability regression. Inference based on a time-series of the predictability coefficient is a more reasonable way of approaching this problem than an unconditional regression.

5 Conclusion

This paper shows that if aggregate consumption and dividends in the economy are not cointegrated and are both driven by a single slow-moving shock that has differential effects in expected consumption and dividend growth, then that has very important ramifications for the PD ratio in the economy with a representative agent with DE preferences. The non-linearities generated in the PD ratio creates a high volatility of returns of the aggregate market portfolio, which translates into high equity premia because the market price of risk due to the long-run risk component is quite substantial. The volatility is not just high but it is stochastic giving rise to time-varying equity premia as well as a predictability relationship. The volatility in that shock necessary to derive other asset pricing characteristics, i.e. high volatility, high equity premia, etc. will make long horizon return coefficients stochastic and sizeable, but also insignificant when that shock is integrated out.

The basis of this result is classical Mertonian mechanics where the value function of the
agent depends on the underlying shock, $X_t$. Due to the simplication that EIS=1, we get a very simple closed form solution of the value function from which we can evaluate the PD ratio directly without resorting to underlying log-linearization that destroys elegant non-linearities that can address key asset pricing questions from a simple setting. A challenge in this setting is to address a cross-section of firms in this economy, obeying the hard constraint that the sum of dividends add up to aggregate dividends. I leave that for future work.
References


6 Appendix

Proof of (2.1.1) The Bellman equation in (4) can be written as

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

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

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

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

Proof of Proposition (2.2.1) 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.
\[
\frac{d \Lambda}{\Lambda} = -r^p_t dt - \gamma \sigma_C dW_C - \frac{\lambda(\gamma - 1)}{\kappa + \beta} \sigma_d dW_X \\
r^p_t = \lambda X_t + \mu_C - \gamma \sigma_C^2 + \beta
\]

**Proof of Proposition (2.2.2)** The stock price of a firm is

\[
P_t = \frac{1}{\Lambda_t} E_t \int_t^\infty \Lambda_s D_s ds
= \frac{1}{\Lambda_t} \int_t^\infty E_t \Lambda_s D_s ds
\]

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, we get the following PDE

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

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

\[
P'_1(\tau) = (1 - \lambda) - \kappa P_1(\tau)
\]
\[
P'_2(\tau) = \mu_D - \mu_C + \gamma \sigma_C^2 - \beta - P_1(\tau)\sigma_x^2 \left[ \frac{\lambda(\gamma - 1)}{\kappa + \beta} - \frac{1}{2} P_1(\tau) \right]
\]

44
with initial conditions $P_1(0) = P_2(0) = 0$. The solution of these ODEs are

\[
P_1(\tau) = \frac{1 - \lambda}{\kappa} (1 - e^{-\kappa \tau}) \\
P_2(\tau) = a \tau + b (e^{-\kappa \tau} - 1) + c (1 - e^{-2\kappa \tau})
\]

\[
a = \mu_D - \mu_C + \gamma \sigma_C^2 - \beta + \frac{\sigma_z^2 (1 - \lambda)}{2 \kappa} \left[ \frac{1 - \lambda}{\kappa} - 2\lambda (\gamma - 1) \right] \\
b = \frac{1 - \lambda}{\kappa} \left[ \frac{\sigma_z^2}{\kappa} \left[ \frac{1 - \lambda}{\kappa} - \frac{\lambda (\gamma - 1)}{\kappa + \beta} \right] \right] \\
c = \frac{\sigma_z^2 (1 - \lambda)^2}{4 \kappa^3}
\]

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

\[
P_t = D_t G(X_t)
\]

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

Cumulative excess return $dR_t = \frac{D_t dt + dp_t dt}{P_t}$ 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 $\mu_t^R = -\text{Cov}_t \left( \frac{d\Lambda_t}{\Lambda_t}, dP_t \right) = \frac{G_X}{G} \frac{\lambda (\gamma - 1)}{\kappa + \beta} \sigma_z^2$.

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

\[
\bar{R}_T = \frac{P_T}{P_t} + \int_t^T D_t dr
\]

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

\[
\mu_G = \frac{\sigma^2}{2} \int_t^\infty \exp(\cdot) P^2_1(\tau) d\tau - \kappa X_t \int_t^\infty \exp(\cdot) P_1(\tau) d\tau - \int_t^\infty \exp(\cdot) (P_1'(\tau) X_t + P_2'(\tau)) d\tau - 1
\]

\[
\sigma_G = \sigma x \int_t^\infty \exp(\cdot) P_1(\tau) d\tau
\]

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

\[
\frac{dP}{P} = \left[ \mu_D + X_t + \frac{\mu G}{G} \right] dt + \sigma_D dW_D + \frac{\sigma G}{G} dW_x
\]

$$= \mu_P(X_t) dt + \sigma_P(X_t) \cdot dW$$

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

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

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

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

Fortunately, the conditional expectation of $\tilde{R}_T$ 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,

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

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

$$\left[ \mu_D + X_t + \frac{\mu G}{G} \right] - \frac{H_x \kappa X_t}{H} + \frac{\sigma_x^2}{2} \frac{H_{XX}}{H} + \frac{H_x \sigma_D \sigma_x}{G} = \frac{H_z}{H}$$

with boundary condition $H(X_t, 0) = 1$. Thus $E_t[P_T] = P_t H(X_t, z)$. Using (21), I can rewrite $E_t[P_T] = P_t \exp \left[ \left( \int_t^T \mu_P(X_s) ds \right) \right]$ which implies $H(X_t, z) = \exp \left( \int_t^T \mu_P(X_s) ds \right)$ 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. $E_t[D_r] = D_t \exp [A(s) X_t + B(s)]$ where

$$A(s; \kappa) = \frac{1 - e^{-\kappa s}}{\kappa}$$

$$B(s) = \mu_D s + \frac{\sigma_x^2}{2 \kappa^2} (s - 2 A(s; \kappa) + A(s; 2 \kappa))$$

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

$$E_t[\tilde{R}_T] = \frac{E_t[P_T + \int_t^T D_r dr]}{P_t} = \frac{E_t[P_T] + \int_t^T E_t[D_r] dr}{P_t}$$

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

$$= \alpha(X_t; T, t) \frac{D_t}{P_t}$$

$H(X_t; T, t) = \exp \left[ \int_t^T \mu_P(X_s) ds \right]$. 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 $\tilde{R}_T$ according to (22) and compute the variance based on simulation.
Figure 1: Time-Series Plot of Aggregate Dividend Growth against Aggregate Consumption Growth.
Figure 2: Histogram of parameter estimates from (18)-(19). I simulate the Gibbs sampler for 25,000 draws from the posterior and discard the first 20,000 for burn-in. These are the histograms from the remaining 5,000 draws for each parameter.
Figure 3: The top graph shows the 2.5-97.5 quintile of the 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 $\mu_C + \lambda X_t$. The bottom graph shows expected dividend growth rate $\mu_D + X_t$ against dividend growth.
The distribution of 6 endogenously determined quantities - instantaneous expected excess return, instantaneous volatility of cumulative return, the dividend yield, volatility of change in price-dividend ratio, the Sharpe ratio and risk-free rate. From the posterior distribution of parameters summarized in Figure 3, the posterior distribution of each of the above quantities is computed by evaluating the quantities for every draw of the parameter $\theta$ and by integrating out the state $X_t$ using its stationary distribution $X_t \sim N \left( \theta, \frac{\sigma^2}{\gamma} \right)$. I use $\beta = 0.01$ and $\gamma = 7.5$. The full distribution of the quantities is presented along with the red line that shows the unconditional average of the quantities in the US postwar data from 1948-2006. All nominal quantities in the data are deflated by CPI.
Figure 5: The top panel plots expected excess return (9) across different growth rates, $X_t$. The middle panel plots the dividend-price ratio, $\frac{1}{G(X_t)}$, across different $X_t$. The bottom panel plots $G_X$, from the instantaneous predictability relationship (11) across growth rates.
Figure 6: The function $H(X_t, z) - 1$ for $z = 1, 3$ for different $X$’s taken from the unconditional distribution of $X_t \sim N\left(0, \frac{\sigma}{\sqrt{2\nu}}\right)$ using the median parameter values in Table 1.
Table 1: The following represents the parameter estimates from estimating the state-space given in (18)-(19) via a Gibbs sampler. The posterior distribution is presented in the form of quantiles from 2.5-th to the 97.5-th quantile of the simulated posterior draws from the Gibbs sampler. The data used for this sample is the difference between annual dividend and consumption (non-durables and services) growth in the US from 1948-2006.

<table>
<thead>
<tr>
<th></th>
<th>0.025</th>
<th>0.25</th>
<th>0.5</th>
<th>0.75</th>
<th>0.975</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_D$</td>
<td>0.0507</td>
<td>0.0566</td>
<td>0.0603</td>
<td>0.0643</td>
<td>0.0737</td>
</tr>
<tr>
<td>$\sigma_C$</td>
<td>0.0085</td>
<td>0.0095</td>
<td>0.0100</td>
<td>0.0107</td>
<td>0.0122</td>
</tr>
<tr>
<td>$\sigma_x$</td>
<td>0.0214</td>
<td>0.0237</td>
<td>0.0250</td>
<td>0.0268</td>
<td>0.0305</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.1636</td>
<td>0.2087</td>
<td>0.2292</td>
<td>0.2486</td>
<td>0.2791</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>0.0489</td>
<td>0.0497</td>
<td>0.0502</td>
<td>0.0506</td>
<td>0.0514</td>
</tr>
<tr>
<td>$\mu_C$</td>
<td>0.0043</td>
<td>0.0078</td>
<td>0.0095</td>
<td>0.0112</td>
<td>0.0144</td>
</tr>
<tr>
<td>$\mu_D$</td>
<td>0.0261</td>
<td>0.0389</td>
<td>0.0460</td>
<td>0.0528</td>
<td>0.0655</td>
</tr>
</tbody>
</table>

Table 2: This Table documents the posterior distribution of market prices of risk given in (6). Given the full parameter distribution summarized in Table 1, 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 $\beta = 0.01$ and $\gamma = 7.5$.

<table>
<thead>
<tr>
<th></th>
<th>0.025</th>
<th>0.25</th>
<th>0.5</th>
<th>0.75</th>
<th>0.975</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transient risk</td>
<td>0.0635</td>
<td>0.0709</td>
<td>0.0753</td>
<td>0.0802</td>
<td>0.0914</td>
</tr>
<tr>
<td>LR risk</td>
<td>0.4278</td>
<td>0.5420</td>
<td>0.6072</td>
<td>0.6705</td>
<td>0.8116</td>
</tr>
</tbody>
</table>
Table 3: The following are sample statistics 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. The data interval is annual from 1948-2006.

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Real Excess Return</td>
<td>0.0744</td>
</tr>
<tr>
<td>Volatility of Real Excess Return</td>
<td>0.1662</td>
</tr>
<tr>
<td>Average Sharpe-Ratio</td>
<td>0.4482</td>
</tr>
<tr>
<td>Average Dividend-Price Ratio</td>
<td>0.0361</td>
</tr>
<tr>
<td>Volatility of the growth in Price-Dividend Ratio</td>
<td>0.1296</td>
</tr>
<tr>
<td>Average Real Risk-free Rate</td>
<td>0.0143</td>
</tr>
</tbody>
</table>

Table 4: Below I present quantiles from the posterior distribution of endogenous quantities with $\beta = 0.01$ and $\gamma = 7.5$. Then using the full parameter distributions obtained through the Gibbs sampler, I compute the posterior distribution of instantaneous expected excess return $- \mu_t^R$ using (9), instantaneous volatility of cumulative return $- \sigma_t^R$ using (2.2.2), Sharpe ratio by dividing $\mu_t^R$ by $\sigma_t^R$, the dividend-price ratio $- \frac{D_t}{P_t}$ by $\frac{1}{G(X)}$ where $G(X)$ is given in (2.2.2), volatility of the growth in price-dividend ratio by $\frac{\sigma_x}{\sigma}$ and the risk-free rate $- r_t^f$ using (7). For all of the following quantities, I integrate out the initial state $X_t$ by using its stationary distribution $X_t \sim N \left(0, \frac{\sigma_x}{\sqrt{2\pi}} \right)$.

<table>
<thead>
<tr>
<th>Quantiles</th>
<th>0.025</th>
<th>0.25</th>
<th>0.5</th>
<th>0.75</th>
<th>0.975</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_t^R$</td>
<td>0.0701</td>
<td>0.0886</td>
<td>0.0955</td>
<td>0.1030</td>
<td>0.1198</td>
</tr>
<tr>
<td>$\sigma_t^R$</td>
<td>0.1430</td>
<td>0.1582</td>
<td>0.1685</td>
<td>0.1830</td>
<td>0.2297</td>
</tr>
<tr>
<td>Sharpe ratio</td>
<td>0.3878</td>
<td>0.5000</td>
<td>0.5585</td>
<td>0.6185</td>
<td>0.7452</td>
</tr>
<tr>
<td>$\frac{D_t}{P_t}$</td>
<td>0.0340</td>
<td>0.0426</td>
<td>0.0550</td>
<td>0.0622</td>
<td>0.0792</td>
</tr>
<tr>
<td>$\text{Vol} \left( \frac{dG}{G} \right)$</td>
<td>0.1296</td>
<td>0.1456</td>
<td>0.1568</td>
<td>0.1720</td>
<td>0.2216</td>
</tr>
<tr>
<td>$r_t^f$</td>
<td>0.0127</td>
<td>0.0170</td>
<td>0.0188</td>
<td>0.0205</td>
<td>0.0237</td>
</tr>
<tr>
<td>Coefficient of Predictability</td>
<td>1.6557</td>
<td>2.0840</td>
<td>2.4137</td>
<td>2.8944</td>
<td>4.9874</td>
</tr>
</tbody>
</table>
Table 5: This table decomposes long-horizon predictability into dividend growth and price growth. I report unconditional quantities of each of the predictability component of dividend growth and PD ratio growth, as well as the $R^2$. I focus on the endogenously determined long-horizon coefficient given by

$$E_t[\bar{R}_T] = \alpha(X_t, z) \frac{D_t}{P_t}$$

where

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

(23)

(24)

In order to simulate the quantities above, I restrict myself to the 0.25, 0.5 and 0.75-th quantiles of parameters given in Table 1. Furthermore, I use $\beta = 0.01$ and $\gamma = 7.5$ and simulate using monthly increment by setting $dt = 1/12$.

For each quantile, I compute the unconditional estimate of $\alpha(z)$ by integrating out $\alpha(X_t, z)$ by using Monte Carlo integration. First, I create a conditional distribution of $H(X_t, z)$, by starting off at different $X_t$'s drawn from $N(0, \frac{1}{\sqrt{2\pi}})$ and simulating forward 10,000 paths of the growth rates from $t$ to $T$ and computing $\exp\left[\int_t^T \mu_p(X_s)ds\right]$ on that path and then taking averages across path to form $H(X_t, z)$. Then, I form an unconditional distribution of $\alpha(\tau)$ by integrating out the above expression of $\alpha(X_t, z)$ by picking $X_t$ again from the unconditional distribution of $X_t$.

The unconditional variance of the conditional mean $E_t[\bar{R}_T]$ as well as the unconditional variance of $\bar{R}_T$ in (15) 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 $\bar{R}_T$ for 10,000 paths 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. I repeat the same exercise to compute the variance of the conditional mean - $Var[E_t[\bar{R}_T]]$. To form $R^2$, I simply divide the variance of the unconditional mean by the variance of $\bar{R}_T$.

Standard errors are in parenthesis.
<table>
<thead>
<tr>
<th>$z(\text{years})$</th>
<th>Quantiles</th>
<th>0.25</th>
<th>0.5</th>
<th>0.75</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\hat{P}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\hat{B}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(35.428)</td>
<td>(29.117)</td>
<td>(24.146)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>$H(1) - 1$</td>
<td>0.056</td>
<td>0.063</td>
<td>0.070</td>
</tr>
<tr>
<td>Div Growth</td>
<td>1.019</td>
<td>1.023</td>
<td>1.027</td>
<td></td>
</tr>
<tr>
<td>$\alpha(1)$</td>
<td>4.325</td>
<td>4.266</td>
<td>4.176</td>
<td></td>
</tr>
<tr>
<td>$Var(\hat{R}_T)$</td>
<td>0.030</td>
<td>0.032</td>
<td>0.036</td>
<td></td>
</tr>
<tr>
<td>$Var[E_t[\hat{R}_T]]$</td>
<td>0.001</td>
<td>0.002</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.037</td>
<td>0.056</td>
<td>0.058</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>$H(2) - 1$</td>
<td>0.126</td>
<td>0.142</td>
<td>0.158</td>
</tr>
<tr>
<td>Div Growth</td>
<td>1.046</td>
<td>1.055</td>
<td>1.063</td>
<td></td>
</tr>
<tr>
<td>$\alpha(2)$</td>
<td>8.625</td>
<td>8.534</td>
<td>8.370</td>
<td></td>
</tr>
<tr>
<td>$Var(\hat{R}_T)$</td>
<td>0.071</td>
<td>0.074</td>
<td>0.112</td>
<td></td>
</tr>
<tr>
<td>$Var[E_t[\hat{R}_T]]$</td>
<td>0.009</td>
<td>0.010</td>
<td>0.141</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.119</td>
<td>0.130</td>
<td>0.125</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>$H(3) - 1$</td>
<td>0.192</td>
<td>0.219</td>
<td>0.245</td>
</tr>
<tr>
<td>Div Growth</td>
<td>1.070</td>
<td>1.083</td>
<td>1.096</td>
<td></td>
</tr>
<tr>
<td>$\alpha(3)$</td>
<td>13.144</td>
<td>12.066</td>
<td>12.870</td>
<td></td>
</tr>
<tr>
<td>$Var(\hat{R}_T)$</td>
<td>0.132</td>
<td>0.135</td>
<td>0.184</td>
<td></td>
</tr>
<tr>
<td>$Var[E_t[\hat{R}_T]]$</td>
<td>0.019</td>
<td>0.026</td>
<td>0.042</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.140</td>
<td>0.193</td>
<td>0.229</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>$H(4) - 1$</td>
<td>0.279</td>
<td>0.320</td>
<td>0.359</td>
</tr>
<tr>
<td>Div Growth</td>
<td>1.105</td>
<td>1.123</td>
<td>1.142</td>
<td></td>
</tr>
<tr>
<td>$\alpha(4)$</td>
<td>18.048</td>
<td>18.078</td>
<td>17.934</td>
<td></td>
</tr>
<tr>
<td>$Var(\hat{R}_T)$</td>
<td>0.217</td>
<td>0.246</td>
<td>0.310</td>
<td></td>
</tr>
<tr>
<td>$Var[E_t[\hat{R}_T]]$</td>
<td>0.038</td>
<td>0.058</td>
<td>0.074</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.185</td>
<td>0.234</td>
<td>0.238</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>$H(5) - 1$</td>
<td>0.355</td>
<td>0.411</td>
<td>0.466</td>
</tr>
<tr>
<td>Div Growth</td>
<td>1.132</td>
<td>1.157</td>
<td>1.182</td>
<td></td>
</tr>
<tr>
<td>$\alpha(5)$</td>
<td>21.956</td>
<td>22.151</td>
<td>24.113</td>
<td></td>
</tr>
<tr>
<td>$Var(\hat{R}_T)$</td>
<td>0.403</td>
<td>0.505</td>
<td>0.635</td>
<td></td>
</tr>
<tr>
<td>$Var[E_t[\hat{R}_T]]$</td>
<td>0.094</td>
<td>0.119</td>
<td>0.157</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.23</td>
<td>0.236</td>
<td>0.246</td>
<td></td>
</tr>
</tbody>
</table>