Risk Aversion, Investor Information, and Stock Market Volatility*

Kevin J. Lansing†
Federal Reserve Bank of San Francisco and Norges Bank

Stephen F. LeRoy‡
UC Santa Barbara and Federal Reserve Bank of San Francisco

June 4, 2012

Abstract

This paper employs a standard asset pricing model with power-utility to derive theoretical volatility measures for the price-dividend ratio and the real equity return in a setting that allows for varying degrees of investor information about future dividends. For reasonable levels of risk aversion, we show that the model can match the observed volatility in long-run U.S. stock market data if investors can accurately foresee future dividends. The variance of the model price-dividend ratio increases monotonically with investor information about future dividends whereas the relationship between equity return variance and information is non-monotonic. We also derive a theoretical variance decomposition for the fundamental price-dividend ratio and show that it differs in important ways from the data. Specifically, even though the model can account for observed stock market volatility, it does so by generating an implausibly volatile risk-free rate combined with an insufficiently predictable excess return on equity.

Keywords: Asset Pricing, Excess Volatility, Variance Bounds, Risk Aversion, Variance Decomposition.

JEL Classification: E44, G12.

* An earlier version of this paper was titled “Risk Aversion and Stock Price Volatility.”
† Corresponding author. Research Department, Federal Reserve Bank of San Francisco, P.O. Box 7702, San Francisco, CA 94120-7702, (415) 974-2393, FAX: (415) 977-4031, email: kevin.j.lansing@sf.frb.org, homepage: www.frbsf.org/economics/economists/klansing.html
‡Department of Economics, University of California, Santa Barbara, CA 93106, (805) 689-2344, email: sleroy@econ.ucsb.edu, homepage: http://www.econ.ucsb.edu/~sleroy/webpage/
1 Introduction

In theory, the price of a stock represents the market’s consensus forecast of the discounted sum of future dividends that will accrue to the owner of the stock. Numerous empirical studies starting with Shiller (1981) and LeRoy and Porter (1981) showed that real-world stock price volatility appeared to vastly exceed the levels implied by the present-value model. A number of econometric problems with the empirical studies were later raised, but it turned out that correcting these problems did not eliminate the appearance of excess volatility.  

Campbell and Shiller (1988) addressed the issue using a log-linear decomposition of the equity return identity. They showed that the observed volatility of stock prices relative to dividends can be attributed mainly to the fact that future equity returns contain a predictable component, contrary to the implications of the constant discount factor model employed in many of the early empirical studies. Predictable future returns imply stochastic discount rates, indicative of risk aversion, whereas the constant discount rate model reflected the assumption of risk neutrality.

Early studies by Grossman and Shiller (1981) and LeRoy and LaCivita (1981) recognized that risk aversion could increase volatility relative to the risk-neutral case. Their arguments, however, were incomplete. Establishing that risk aversion may affect volatility does not, by itself, have implications for the presence or absence of excess volatility. This is so because risk aversion also affects the upper-bound volatility measure computed from “perfect foresight” (or “ex post rational”) stock prices. Consequently, while high risk aversion may imply high volatility, it may or may not imply excess volatility.

This paper, like Campbell and Shiller (1988), makes use of log-linear methods to examine the connections between discounts rates and stock market volatility. However, rather than log-linearizing the return identity, we compute a log-linear approximation of the representative investor’s first-order condition. This approximation incorporates a model specification for the stochastic discount factor and a process for consumption/dividends. Given the model’s variance predictions, we are then able to map our results into the Campbell-Shiller decomposition framework, as discussed further below.

A fact that is often glossed over in discussions of stock market efficiency is that the proposition being tested is a compound null hypothesis. Stock prices are taken to equal the present value of future dividends with the univariate process for dividends taken as given. However, the null hypothesis is silent about how much information investors condition on when forming their expectations of future dividends. LeRoy and Porter (1981) thought of dividends as being generated jointly with other variables by a multivariate ARMA process, which leaves open the possibility that other variables (current earnings, for example) could serve as predictors.

---

1For summaries of this extensive literature, see West (1988a), Gilles and LeRoy (1991), Shiller (2003), and LeRoy (2010).
of future dividends. We show that the existence of such auxiliary information has important quantitative implications for volatility. Interestingly, recent research on business cycles has focused on “news shocks” as an important quantitative source of economic fluctuations. In these models, news shocks provide agents with information about future fundamentals, i.e., future technology innovations.\(^2\)

We begin our analysis with a simple example involving risk-neutral investors and stationary dividends, reminiscent of the assumptions employed in the early empirical studies on excess volatility. Within this simple setting, we show that increasing investor information about future dividend innovations will monotonically raise the unconditional variance of stock prices but monotonically lower the unconditional variance of the “excess payoff,” which measures the return on the stock investment relative to the risk-free asset. The upper bound on unconditional price variance coincides with the lower bound (of zero) on excess payoff variance. These bounds are reached simultaneously when investors are assumed to have perfect foresight about the entire future path of dividends. The upper bound on excess payoff variance is reached when investors are assumed to have no auxiliary information about future dividend innovations, i.e., when investors only know current and past dividends—the typical information assumption employed in asset pricing models. Since the unconditional variance of excess payoffs is nearly the same as the conditional variance of prices, the variance bounds test is effectively a joint test on the unconditional and conditional variances implied by the present-value model.

Our simple example shows that variance bounds for asset pricing variables can be generated by imposing extreme hypothetical specifications for investors’ information about future dividends. Given the explicit focus on information assumptions, variance-bounds tests remain a useful analytical tool for assessing the success or failure of a particular asset pricing model.

When economists talk about stock price volatility it is not clear whether unconditional or conditional variance best corresponds to what they have in mind. LeRoy (1984) and Kleidon (1986) used numerical examples to illustrate the idea that different conclusions may be drawn from considering conditional variances rather than unconditional variances. The frequent use of terms like “choppiness” or “smoothness” in describing observed stock prices suggests that the conditional variance is the appropriate concept, since these terms are taken to refer to short-term price volatility. The fact that unconditional and conditional variances can be pushed in opposite directions by changes in investor information is an important insight derived from the variance bounds tests.

This paper compares volatility measures computed from long-run U.S. stock market data to model-predicted volatility measures in a setting that allows for risk aversion and varying degrees of investor information about future dividends. Using variance bounds tests based on the price-dividend ratio (a stationary variable), we find evidence of excess volatility in long-

\(^2\)See, for example, Barsky and Sims (2011).
run U.S. data for risk aversion coefficients below about 5. For higher degrees of risk aversion, we find that volatility is not excessive if we assume that investors can accurately predict dividends into the distant future. However, to the extent that this information assumption is viewed as implausible, it follows that observed volatility is excessive in that case as well.

In settings with exponentially-growing dividends (as here), the log return variance is the analog to the arithmetic price-change variance measures examined by West (1988b) and Engel (2005) in risk-neutral settings with linearly-growing dividends. These authors showed that the arithmetic price-change variance is a monotonically decreasing function of investors’ information about future dividends. We show by counterexample that when investors are risk averse, the log return variance analogs to the West-Engel results do not go through; log return variance is not a monotonic decreasing function of investors’ information about future dividends. Therefore, imposing extreme hypothetical specifications for investors’ information does not establish bounds on equity return variance within our standard model. We demonstrate that, despite the absence of theoretical variance bounds for log returns, the power-utility model can actually match the observed volatility of log stock returns in long-run U.S. data for risk-aversion coefficients around 4, provided that investors possess some auxiliary information about future dividend innovations.

The intuition for the ambiguous relationship between investor information and return variance is linked to the discounting mechanism. Two crucial elements are the persistence of dividend growth and the investor’s discount factor (which depends on risk aversion). Both elements affect the degree to which future dividend innovations influence the perfect foresight price via discounting from the future to the present. When dividends are sufficiently persistent and the investor’s discount factor is sufficiently close to unity, the discounting weights applied to successive future dividend innovations decay gradually. Since log returns are nearly the same as log price-changes, computation of the log return under perfect foresight tends to “difference out” the future dividend innovations, thus shrinking the magnitude of the perfect foresight return variance relative to the case where the investor has no information about future dividend innovations. In contrast, when dividend growth is less persistent and/or the investor’s discount factor is much less than unity, the discounting weights applied to successive future dividend innovations decay rapidly. Consequently, these terms do not tend to difference out which serves to magnify the perfect foresight return variance relative to the case where the investor has no information about future dividend innovations.

It is straightforward to extend the analysis of return variance to consider the effects of investor information on the variance of excess returns, i.e., the variance of the equity premium. In this case, we can establish a theoretical lower bound (of zero) on the variance of excess log returns when risk averse investors have perfect foresight. This result is directly analogous

---

3On page 41, West (1988b) acknowledges that his result “may not extend immediately if logarithms or logarithmic differences are required to induce stationarity” in dividends.
to our simple example (described above) where the variance of excess payoffs is zero under
the joint assumptions of risk neutrality and perfect foresight. However, when moving from
no information about future dividends to some information (i.e., allowing investors to see
dividends one period ahead), the presence of risk aversion creates an ambiguous relationship
between information and the variance of excess log returns. Hence, we cannot establish a
theoretical upper bound on the variance of excess log returns in the presence of risk aversion.
We note that our results on this topic are similar in some respects to those derived by Veronesi
(2000), even though the concept of information in the two setups is quite different.

Last, we derive a theoretical variance decomposition for the model price-dividend ratio
under different information sets. Much of the previous work in this area is empirical. Specifi-
cally, we examine the degree to which movements in the price-dividend ratio can be accounted
for by movements in either: (1) future dividend growth rates, (2) future risk-free rates, or
(3) future excess returns on equity. In this way, we are able to map our theoretical results to
the empirical findings of Campbell and Shiller (1988), Campbell (1991), and Cochrane (1992,
2005, 2008). We believe we are the first to show that the theoretical variance decomposition
in the power-utility model depends crucially (and almost exclusively) on the risk aversion
coefficient. As risk aversion increases, the representative investor’s stochastic discount factor
becomes more volatile, which in turns raises the variance contribution from future risk-free
rates and lowers the contribution from future dividend growth rates.

The variance contribution from future excess returns in the power-utility model turns out
to be exactly zero, or close to zero, depending on the information set. This is so because future
excess returns in the model are generally not predictable using the current price-dividend ratio.
In contrast, the empirical decomposition shows that the bulk of the variance in the observed
price-dividend ratio is attributable to future excess returns on equity, in stark contrast to the
model’s predictions.

Some recent contributions that allow for bubbles or employ models with time-varying
risk aversion or time-varying volatility of consumption growth have achieved more success in
matching the empirical variance decomposition because these features introduce persistence
(or predictability) in the law of motion for excess returns. For example, Engsted et al. (2012)
introduce a periodically-collapsing rational bubble into a simple asset pricing model with
risk-neutral investors. They show that predictability regressions on simulated data from the
model yield results that are very similar to those obtained by Cochrane (2008) using U.S. data.
Models by Campbell and Cochrane (1999) and Bansal and Yaron (2004) are the most notable
examples that employ time-varying risk aversion and time-varying volatility of consumption
growth, respectively. However, these models must still rely on the assumption of very high
risk aversion to match features of the data. Overall, we conclude that fundamental models
with reasonable levels of risk aversion cannot account for important aspects of real-world stock
prices.
The remainder of the paper is organized as follows. Section 2 derives theoretical variance bounds in a simple setting with risk-neutral investors and stationary dividends. Sections 3 through 6 extend the analysis to a standard power-utility model with exponentially-growing dividends. Section 7 maps our theoretical results to the empirical framework used by Campbell and Shiller (1988) and others. Section 8 concludes. An appendix provides the details for all derivations.

2 Simple Example: Variance Bounds with Risk-Neutral Investors and Stationary Dividends

This section briefly reviews some variance bounds that obtain in a simple setting where investors are risk neutral and dividends are generated by a stationary linear process. The absence of arbitrage implies that the equilibrium stock price \( p_t \) obeys
\[
p_t = E(p_t^* | I_t),
\]
where \( I_t \) represents investors’ information about future dividend realizations and \( p_t^* \) is the perfect foresight price. As originally set forth in Shiller (1981) and LeRoy and Porter (1981), the fact that \( p_t = E(p_t^* | I_t) \) implies that the variance of \( p_t \) is an upper bound on the variance of \( p_t^* \). As LeRoy and Porter (1981) showed, we can also establish a lower bound on the variance of \( p_t \). Define \( H_t = \{d_t, d_{t-1}, d_{t-2}, \ldots \} \) as the information set consisting only of current and past dividends, and define \( \hat{p}_t = E(p_t^* | H_t) \), where \( \hat{p}_t \) is the appropriate stock price for an econometrician who has no information useful in forecasting \( p_t \) other than current and past dividends.\(^4\)

Suppose that investors’ information \( I_t \) contains at least \( H_t \), so that \( H_t \subseteq I_t \), but investors may also have auxiliary information over and above current and past dividends that is useful in predicting \( p_t^* \). For example, forecasts about future earnings combined with historical dividend payout ratios are likely to help predict future dividends even given current and past dividends. The simplest characterization of this idea (to be employed below) defines \( J_t = H_t \cup d_{t+1} \), so that investors can see dividends without error one period ahead. Thus, \( J_t \) is a specific example of \( I_t \) that is intermediate between \( H_t \) and perfect information about the future. The fact that \( J_t \) is a refinement of \( H_t \) implies that \( \text{Var}(\hat{p}_t) \) is a lower bound for \( \text{Var}(\bar{p}_t) \), where \( \bar{p}_t \equiv E(p_t^* | J_t) \). Thus we have
\[
\text{Var}(\hat{p}_t) \leq \text{Var}(\bar{p}_t) \leq \text{Var}(p_t^*),
\]
where the upper and lower variance bounds can be calculated from a univariate model for dividends. The theoretical variance bounds can thus be derived without explicitly specifying the extent of investors’ auxiliary information.

\(^4\)Throughout the paper, we adopt the notation of using stars “*” to denote perfect foresight variables, hats “~” to denote variables computed using information set \( H_t \), overbars “—” to denote variables computed using information set \( J_t = H_t \cup d_{t+1} \), and unmarked variables (such as \( p_t \)) to denote variables computed using the unspecified information set \( I_t \).
We define the “excess payoff” under information set \( I_t \) as

\[
\nu_{t+1} \equiv p_{t+1} + d_{t+1} - \beta^{-1} p_t,
\]

which represents next-period’s cash value from the equity investment minus the payoff from an equal investment in the risk-free asset. Under the risk-neutral utility function \( \sum_{t=0}^{\infty} \beta^t c_t \), where \( \beta \in (0, 1) \) is the subjective time discount factor and \( c_t \) is consumption, it is straightforward to show that the gross risk-free rate equals \( \beta^{-1} \). Hence, \( \nu_{t+1}/p_t \) represents the excess return on equity relative to the risk-free asset. From the investor’s first-order condition for equity holdings, we have \( p_t = \beta E \{ (p_{t+1} + d_{t+1}) | I_t \} \), which implies that the excess payoff (2) is simply the one-period-ahead forecast error, which is iid over time under all information specifications. Multiplying successive iterations of equation (2) by \( \beta^i \) for \( i = 1, 2, 3, \ldots \) and then summing across the resulting equations yields

\[
\beta \nu_{t+1} + \beta^2 \nu_{t+2} + \beta^3 \nu_{t+3} + \ldots = -p_t + \beta d_{t+1} + \beta^2 d_{t+2} + \beta^3 d_{t+3} + \ldots.
\]

Solving equation (3) for \( p_t^* \) and then taking the variance of both sides yields

\[
Var(p_t^*) = Var(p_t) + \frac{\beta^2}{1-\beta^2} Var(\nu_t),
\]

where we have assumed that dividends are generated by a stationary linear process so that the variances are constant. The perfect-foresight version of the first-order condition is \( p_t^* = \beta (p_{t+1} + d_{t+1}) \), which shows that the excess payoff under perfect foresight is zero for all \( t \) such that \( Var(\nu_t^*) = 0 \). Since \( Var(p_t^*) - Var(\nu_t^*) \geq 0 \) from equation (1), the above expression establishes that \( Var(\nu_t) \geq 0 = Var(\nu_t^*) \).

Similarly, we define the excess payoff under information set \( H_t \) as \( \tilde{\nu}_{t+1} \equiv \tilde{p}_{t+1} + d_{t+1} - \beta^{-1} \tilde{p}_t \). Following the same methodology as above, we obtain

\[
Var(p_t^*) = Var(\tilde{p}_t) + \frac{\beta^2}{1-\beta^2} Var(\tilde{\nu}_t).
\]

Substituting for \( Var(p_t^*) \) from equation (4) into equation (5) and noting that \( Var(\tilde{p}_t) - Var(\tilde{\nu}_t) \geq 0 \) from equation (1) establishes that \( Var(\tilde{\nu}_t) \geq Var(\nu_t) \). Thus, if investors are risk neutral and dividends are generated by a stationary linear process, then we have the following bounds on excess payoff variance previously derived in LeRoy (1996):

\[
Var(\nu_t^*) = 0 \leq Var(\nu_t) \leq Var(\tilde{\nu}_t).\]

In the above example, the more information agents have about future dividend innovations, the higher is the variance of prices and lower is the variance of excess payoffs. The maintained lower bound on investors’ information is represented by \( H_t \). The excess payoff variance associated with \( H_t \) therefore represents an upper bound for the excess payoff variance associated with \( I_t \).
3 Allowing for Risk Aversion and Growing Dividends

We now extend the analysis to a more realistic environment with risk-averse investors and exponentially-growing dividends. Equity shares are priced as in the frictionless pure exchange model of Lucas (1978). A representative investor can purchase shares to transfer wealth from one period to another. Each share pays an exogenous stream of stochastic dividends in perpetuity. The representative investor’s problem is to maximize

\[ E \left\{ \sum_{t=0}^{\infty} \beta^t \frac{c_1^{1-\alpha} - 1}{1-\alpha} |I_0| \right\} , \]  

subject to the budget constraint

\[ c_t + p_t s_t = (p_t + d_t) s_{t-1}, \quad c_t, \ s_t > 0, \]  

where \( c_t \) is the investor’s consumption in period \( t \), \( \alpha \) is the coefficient of relative risk aversion and \( s_t \) is the number of shares held in period \( t \). The first-order condition that governs the investor’s share holdings is

\[ p_t = E \left\{ \beta \left( \frac{c_{t+1}}{c_t} \right)^{-\alpha} (p_{t+1} + d_{t+1}) |I_t| \right\} . \]  

The first-order condition can be iterated forward to substitute out \( p_{t+j} \) for \( j = 1, 2, ... \) Applying the law of iterated expectations and imposing a transversality condition that excludes bubble solutions yields the following expression for the equilibrium stock price:

\[ p_t = E \left\{ \sum_{j=1}^{\infty} M_{t,t+j} d_{t+j} |I_t| \right\} , \]  

where \( M_{t,t+j} = \beta^j (c_{t+j}/c_t)^{-\alpha} \) is the stochastic discount factor. The perfect foresight price is given by

\[ p^*_t = \sum_{j=1}^{\infty} M_{t,t+j} d_{t+j} . \]  

Equity shares are assumed to exist in unit net supply. Market clearing therefore implies \( c_t = d_t \) for all \( t \).

We assume that the growth rate of dividends \( x_t \equiv \log (d_t/d_{t-1}) \) is governed by the following AR(1) process:

\[ x_{t+1} = \rho x_t + (1 - \rho) \mu + \varepsilon_{t+1}, \quad \varepsilon_{t+j} \sim N (0, \sigma^2_\varepsilon) , \text{ iid} , \quad |\rho| < 1 . \]  

7
In the special case of $\rho = 0$, the above specification implies that the level of real dividends follows a geometric random walk with drift, as in LeRoy and Parke (1992). Our specification implies that the unconditional moments of consumption/dividend growth are given by

$$E(x_t) = \mu,$$

$$Var(x_t) = \frac{\sigma_x^2}{1 - \rho^2},$$

$$Cov(x_{t+j}, x_t) = \rho^j Var(x_t).$$

### 4 Volatility of the Price-Dividend Ratio

Since dividends and equilibrium stock prices trend upward, variance measures conditional on some initial date will increase with time. To avoid this time-varying volatility result, a trend correction must be imposed. The solution adopted by Shiller (1981) was to assume that dividends and prices are stationary around a time trend. In the presence of a unit root, specifying reversion to a time trend leads to a downward-biased volatility estimate for the variable in question. Moreover, the trend specification employed by Shiller is not realistic for some variables and sample periods; mean-reversion to a time trend induces negative autocorrelation in growth rates, which conflicts with what we see in U.S. data for the growth rates of real dividends, real stock prices, and post-World War II real per capita consumption. We note, however, that data on real per capita consumption over the period 1890 to 2008 does exhibit weak negative autocorrelation in growth rates.

LeRoy and Porter (1981) corrected for nonstationarity by reversing the effect of earnings retention on dividends and stock prices, but that procedure appeared to produce series that were not stationary. Current practice is to correct for trend by working with intensive variables, such as the price-dividend ratio or the rate of return, as these variables will be stationary in the models of interest (see, for example, Cochrane 1992 and LeRoy and Parke 1992).

The price-dividend ratios implied by the information sets $H_t$ and $J_t$ are denoted by $\hat{y}_t \equiv \hat{p}_t/d_t$ and $\overline{y}_t \equiv \overline{p}_t/d_t$, respectively, while the perfect foresight price-dividend ratio is denoted by $y^*_t \equiv p^*_t/d_t$. By substituting the equilibrium condition $c_t = d_t$ into the first-order condition

---

5West (1988a, p. 641) summarizes the various assumptions made in the literature regarding the stochastic process for dividends and prices.

(9), the first-order condition under the various information assumptions can be written as

\[
\hat{y}_t = E \left\{ \beta \exp \left[ (1 - \alpha) x_{t+1} \right] (\hat{y}_{t+1} + 1) | H_t \right\}, \tag{16}
\]

\[
\bar{y}_t = E \left[ \beta \exp \left[ (1 - \alpha) x_{t+1} \right] (\bar{y}_{t+1} + 1) | J_t \right], \tag{17}
\]

\[
y_t^* = \beta \exp \left[ (1 - \alpha) x_{t+1} \right] (y_{t+1}^* + 1). \tag{18}
\]

The fact that \( \hat{y}_t \) and \( y_t^* \) are ratios with the same denominator \( d_t \), together with the fact that \( d_t \) is measurable under all information specifications, immediately implies

\[
\text{Var}(\hat{y}_t) \leq \text{Var}(\bar{y}_t) \leq \text{Var}(y_t^*). \tag{19}
\]

Hence, the basic form of the variance bound derived in the earlier literature under risk neutrality, i.e., \( \text{Var}(\hat{p}_t) \leq \text{Var}(\bar{p}_t) \leq \text{Var}(p_t^*) \), carries over to the case of risk aversion when the price-dividend ratio (an intensive variable) is substituted for the stock price (an extensive variable).

4.1 Variance under Information Set \( H_t \)

We next derive an approximate analytical solution for the variance of \( \hat{y}_t \) under information set \( H_t \). This involves solving the first-order condition (16) subject to the dividend growth process (12). To do so, it is convenient to define the following nonlinear change of variables:

\[
\tilde{z}_t \equiv \beta \exp \left[ (1 - \alpha) x_t \right] (\hat{y}_t + 1), \tag{20}
\]

where \( \tilde{z}_t \) represents a composite variable that depends on both the growth rate of dividends and the price-dividend ratio. The first-order condition (16) becomes

\[
\hat{y}_t = E(\tilde{z}_{t+1}|H_t), \tag{21}
\]

implying that \( \hat{y}_t \) is simply the rational forecast of the composite variable \( \tilde{z}_{t+1} \), conditioned on \( H_t \). Combining (20) and (21), the composite variable \( \tilde{z}_t \) is seen to be governed by the following equilibrium condition:

\[
\tilde{z}_t = \beta \exp \left[ (1 - \alpha) x_t \right] [E(\tilde{z}_{t+1}|H_t) + 1], \tag{22}
\]

which shows that the value of \( \tilde{z}_t \) in period \( t \) depends on the conditional forecast of that same variable in period \( t + 1 \).

The following proposition presents an approximate analytical solution for the composite variable \( \tilde{z}_t \).
**Proposition 1.** An approximate analytical solution for the equilibrium value of the composite variable $\tilde{z}_t$ under information set $H_t$ is given by

$$ \tilde{z}_t = a_0 \exp \left[ a_1 (x_t - \mu) \right], $$

where $a_1$ solves

$$ a_1 = \frac{1 - \alpha}{1 - \rho \beta \exp \left[ (1 - \alpha)\mu + \frac{1}{2} (a_1)^2 \sigma^2 \varepsilon \right]} \quad \text{and} \quad a_0 \equiv \exp \{ E[\log (\tilde{z}_t)] \} \text{ is given by} $$

$$ a_0 = \frac{\beta \exp \left[ (1 - \alpha)\mu \right]}{1 - \beta \exp \left[ (1 - \alpha)\mu + \frac{1}{2} (a_1)^2 \sigma^2 \varepsilon \right]}, $$

provided that $\beta \exp \left[ (1 - \alpha)\mu + \frac{1}{2} (a_1)^2 \sigma^2 \varepsilon \right] < 1.$

**Proof:** See Appendix A.1.

Two values of $a_1$ satisfy the nonlinear equation in Proposition 1. The inequality restriction selects the value of $a_1$ with lower magnitude to ensure that $a_0$ is positive.\(^7\) Given the approximate solution for the composite variable $\tilde{z}_t$, we can recover $\tilde{y}_t$ as follows:

$$ \tilde{y}_t = E(\tilde{z}_{t+1}|H_t) \approx a_0 \exp \left[ a_1 \rho (x_t - \mu) + \frac{1}{2} (a_1)^2 \sigma^2 \varepsilon \right]. \quad (23) $$

As shown in Appendix A.2, the approximate fundamental solution can be used to derive the following unconditional variance of the log price-dividend ratio:

$$ \text{Var} [\log (\tilde{y}_t)] = (a_1 \rho)^2 \text{Var} (x_t), \quad (24) $$

which in turn can be used to derive an expression for $\text{Var} (\tilde{y}_t).$\(^8\)

From equation (23), the direction of the effect of dividend growth fluctuations on $\tilde{y}_t$ depends on the sign of the product $a_1 \rho$. Suppose first that $\rho < 0$, so that agents expect that high current dividend growth will be followed by low growth. Assuming $\alpha < 1$ such that $a_1 > 0$, we have $a_1 \rho < 0$ which causes stocks to trade at a lower-than-average multiple of current dividends today, i.e., a lower value of $\tilde{y}_t$, if current dividend growth is high. On the other hand when $\alpha > 1$ such that $a_1 < 0$, we have $a_1 \rho > 0$. In this case, the expected lower lower

\(^7\)Lansing (2010) compares the approximate solution from Proposition 1 to the exact theoretical solution derived by Burnside (1998). The approximate solution is extremely accurate for low and moderate levels of risk aversion ($\alpha \approx 2$). But even for high levels of risk aversion ($\alpha \approx 10$), the approximation error for the equilibrium price-dividend ratio remains below 5 percent.

\(^8\)Given the unconditional mean $E[\log (\tilde{y}_t)] = \log (a_0) + (a_1)^2 \sigma^2 / 2$ and the expression for $\text{Var} [\log (\tilde{y}_t)]$ from equation (24), the unconditional variance of $\tilde{y}_t$ itself can be computed by making use of the following expressions for the mean and variance of the log-normal distribution: $E (\tilde{y}_t) = \exp \{ E[\log (\tilde{y}_t)] + \frac{1}{2} \text{Var} [\log (\tilde{y}_t)] \}$ and $\text{Var} (\tilde{y}_t) = E (\tilde{y}_t)^2 \{ \exp \{ \text{Var} [\log (\tilde{y}_t)] \} - 1 \}$. 

10
dividend growth in the following period is more than offset by a high realization of the stochastic discount factor, leading to a higher value of \( \hat{y}_t \) today. All of these effects are reversed when \( \rho > 0 \).

In the special case of logarithmic utility we have \( \alpha = 1 \) such that \( a_1 = 0 \). In this case, fluctuations in dividend growth do not affect \( \log ( \hat{y}_t ) \), which is therefore constant. This is because the income and substitution effects of a shock to dividend growth are exactly offsetting. From equation (24), it is easy to see intuitively how different levels of \( a \) affect the variance of \( \log ( \hat{y}_t ) \):

When \( \alpha < 1 \), an increase in \( \alpha \) shrinks the magnitude of \( a_1 \) which moves the variance of \( \log ( \hat{y}_t ) \) toward zero. This happens because fluctuations in dividend growth are increasingly offset by fluctuations in their marginal utility; the closer \( \alpha \) is to unity, the greater is the offset. When \( \alpha > 1 \), an increase in \( \alpha \) raises the magnitude of \( a_1 \). In this case, higher risk aversion raises the extent to which fluctuations in marginal utility exceed fluctuations in inverse consumption, thereby increasing the variance of \( \log ( \hat{y}_t ) \).

4.2 Variance under Information Set \( J_t = H_t \cup d_{t+1} \)

In the preceding subsection we assumed that investors have no auxiliary information that would help to predict future dividends. We now relax that assumption by allowing investors to see dividends one period ahead without error, as in LeRoy and Parke (1992). This setup seems particularly realistic in light of company-provided guidance about future financial performance which is typically disseminated to the public via quarterly conference calls. The expanded information set is defined as \( J_t = H_t \cup d_{t+1} = \{ d_{t+1}, d_t, d_{t-1}, d_{t-2}, \ldots \} \). The set \( J_t \) is an example of an investor information set that is strictly finer than \( H_t \) but strictly coarser than the perfect information underlying \( p_t^* \).

As shown in Appendix B.1, the expanded information set \( J_t \) implies the following relationships:

\[
\begin{align*}
\bar{p}_t &= M_{t,t+1} (d_{t+1} + \hat{p}_{t+1}) , \\
\bar{y}_t &= \beta \exp \left[(1 - \alpha) x_{t+1}\right] (1 + \hat{y}_{t+1}) \\
&= \hat{z}_{t+1} .
\end{align*}
\]

As specified above, \( \bar{p}_t \) and \( \bar{y}_t \) are the price and price-dividend ratio under \( J_t \), while \( \hat{p}_t \) and \( \hat{y}_t \) are their counterparts under \( H_t \). Under information set \( J_t \), the discount factor \( M_{t,t+1} \) is known to investors at time \( t \). From equations (21) and (26), it follows directly that \( \hat{y}_t = E(\bar{y}_t | H_t) \), which in turn implies \( Var (\hat{y}_t) \leq Var (\bar{y}_t) \).

From equations (26) and Proposition 1, the approximate law of motion for \( \bar{y}_t = \hat{z}_{t+1} \) implies the following unconditional variance:

\[
Var [\log (\bar{y}_t)] = (a_1)^2 Var (x_t) .
\]
Comparing the above expression to the expression for $\text{Var} \left[ \log \left( \frac{b}{y_t} \right) \right]$ from equation (24) shows that $\text{Var} \left[ \log \left( \frac{b}{y_t} \right) \right] \leq \text{Var} \left[ \log (y_t) \right]$ since $|\rho| < 1$.

### 4.3 Variance under Perfect Foresight

The assumption of perfect foresight represents an upper bound on investors’ information about future dividends. The perfect foresight price-dividend ratio $y_t^*$ is governed by equation (18), which is a nonlinear law of motion. To derive an analytical expression for the perfect foresight variance, we approximate equation (18) using the following log-linear law of motion (Appendix C.1):

$$
\log \left( y_t^* \right) - E \left[ \log \left( y_t^* \right) \right] \simeq (1-\alpha) (x_{t+1} - \mu) + \beta \exp \left[ \left( 1-\alpha \right) \mu \right] \left\{ \log \left( y_{t+1}^* \right) - E \left[ \log \left( y_t^* \right) \right] \right\}.
$$

(28)

The approximate law of motion (28) and the dividend growth process (12) can be used to derive the following unconditional variance (Appendix C.2):

$$
\text{Var} \left[ \log \left( y_t^* \right) \right] = \frac{(1-\alpha)^2}{1 - \beta^2 \exp \left[ 2(1-\alpha)\mu \right]} \left\{ \frac{1 + \rho \beta \exp \left[ (1-\alpha)\mu \right]}{1 - \rho \beta \exp \left[ (1-\alpha)\mu \right]} \right\} \text{Var} \left( x_t \right),
$$

(29)

which is considerably more complicated than either $\text{Var} \left[ \log \left( \frac{\tilde{y}_t}{y_t} \right) \right]$ from equation (24) or $\text{Var} \left[ \log (\bar{y}_t) \right]$ from equation (27).

### 4.4 Model Calibration

Given that the Lucas model implies $c_t = d_t$ in equilibrium, we calibrate the stochastic process for $x_t$ in equation (12) using U.S. annual data for the growth rate of per capita real consumption from 1890 to 2008.\(^9\) We choose parameters to match the mean, standard deviation, and autocorrelation of per capita consumption growth in the data using the moment formulas given by equations (13) through (15). This procedure yields $\mu = 0.0203$, $\sigma_\epsilon = 0.0351$, and $\rho = -0.1$. For each value of $\alpha$, we calibrate the subjective time discount factor $\beta$ so as to achieve $E \left[ \log \left( \frac{\tilde{y}_t}{y_t} \right) \right] = 3.18$ in the model, consistent with the sample average value of the log price-dividend ratio for the S&P 500 stock index from 1871 to 2008.\(^10\) When $\alpha$ exceeds a value of about 3, achieving the target value of $E \left[ \log \left( \frac{\tilde{y}_t}{y_t} \right) \right]$ in the model requires a value of $\beta$ that is greater than unity. Nevertheless, for all values of $\alpha$ examined, the mean value of the stochastic discount factor $E \left[ \beta \left( c_{t+1}/c_t \right)^{-\alpha} \right]$ remains below unity.\(^11\)


\(^10\)Cochrane (1992) employs a similar calibration procedure. For a given discount factor $\beta$, he chooses the risk aversion coefficient $\alpha$ to match the mean price-dividend ratio in the data.

\(^11\)Kocherlakota (1990) shows that a well-defined competitive equilibrium with positive interest rates can still exist in growth economies when $\beta > 1$. 

12
4.5 Quantitative Analysis

The top panel of Figure 1 compares the variance of the log price-dividend ratio for the S&P 500 index (cross-hatched green line) with the model-computed volatilities for \( \log (\tilde{y}_t) \) (solid blue line), \( \log (\bar{y}_t) \) (dotted grey line), and \( \log (y^*_t) \) (dashed red line).

The standard deviation of \( \log (\tilde{y}_t) \) is close to zero for all values of \( \alpha \). This low figure reflects the fact that the calibrated autocorrelation of dividend growth \( \rho = -0.1 \) is close to zero, corresponding to a near-geometric random walk in the level of dividends.\(^{12}\) The model-predicted volatility for \( \log (\tilde{y}_t) \) is much lower than the standard deviation of the log price-dividend ratio in U.S. data for the period 1871 to 2008, which is 0.41.\(^{13}\) The model-predicted volatility for \( \log (\bar{y}_t) \), which is based on the assumption that investors see dividends one period ahead, is noticeably higher than the volatility of \( \log (\tilde{y}_t) \), but still well below the value observed in the data. These findings suggest the presence of excess volatility in the data, but do not conclusively demonstrate its existence because real-world investors may possess additional information about future dividend growth innovations which would serve to increase the volatility of the U.S. price-dividend ratio.

According to the variance bounds, a finding of excess volatility requires the variance in U.S. data to exceed \( \text{Var} [\log (y^*_t)] \). Figure 1 shows that excess volatility prevails for \( \alpha < 5 \). In contrast, for \( \alpha > 5 \), the variance in U.S. data is below the upper bound of \( \text{Var} [\log (y^*_t)] \), so we cannot make a definitive finding of excess volatility. The interpretation is that the volatility of the log price-dividend ratio in U.S. data is consistent with the model if real-world investors are risk averse and have access to very good information about future dividends. The finding that the theoretical variance inequality is not satisfied when risk aversion is low is consistent with the early variance-bounds tests, which found excess volatility under the assumption of risk neutrality. A conclusion that observed volatility in the data is excessive depends on whether risk aversion coefficients around 5 can be viewed as realistic (most empirical estimates are more like 2), and also on whether it is reasonable to assume that investors can predict dividends into the distant future.

5 Return Volatility

We observed in the introduction that notions of stock market volatility can be connected either with unconditional variance measures, corresponding to a long-run concept of volatility, or with conditional variance measures, corresponding to the short-run. We also noted that, based on earlier research assuming risk neutrality, the present-value model has implications for both measures of volatility. Specifically, the variance-bounds tests involve determining whether

\(^{12}\)Recall that when \( \rho = 0 \), we have \( \text{Var} [\log (\tilde{y}_t)] = 0 \) from equation (24).

\(^{13}\)The standard deviation of the U.S. price dividend ratio in levels (as opposed to logarithms) is 13.8, with a corresponding mean value of 26.6.
the joint restrictions implied by the present-value model for unconditional and conditional volatility measures are satisfied. So far we have concentrated on bounds for unconditional volatility as embodied in \( \text{Var} \{ \log (\hat{y}_t) \} \) and \( \text{Var} \{ \log (y^*_t) \} \).

There are several ways to gauge short-run volatility, including the variance the log price change or the variance of the log equity return. Since these measures are highly correlated, it does not matter much for the substantive results which measure is used.\(^{14}\) Here we examine the model’s implications for the variance of log equity returns under different information sets.

The gross rates of return on equity under each information set can be written as

\[
\hat{R}_{t+1} = \frac{\hat{p}_{t+1} + d_{t+1}}{\hat{p}_t} = \exp (x_{t+1}) \left( \frac{\hat{y}_{t+1} + 1}{\hat{y}_t} \right)
\]

\[
= \beta^{-1} \exp (\alpha x_{t+1}) \left[ \frac{\hat{z}_{t+1}}{E(\hat{z}_{t+1}|H_t)} \right], \quad (30)
\]

\[
\overline{R}_{t+1} = \frac{\overline{p}_{t+1} + d_{t+1}}{\overline{p}_t} = \exp (x_{t+1}) \left( \frac{\overline{y}_{t+1} + 1}{\overline{y}_t} \right)
\]

\[
= \beta^{-1} \exp (\alpha x_{t+1}) \left[ \frac{\overline{z}_{t+1}}{E(\overline{z}_{t+1}|J_t)} \right], \quad (31)
\]

\[
R^*_{t+1} = \frac{p^*_{t+1} + d_{t+1}}{p^*_t} = \exp (x_{t+1}) \left( \frac{y^*_{t+1} + 1}{y^*_t} \right)
\]

\[
= \beta^{-1} \exp (\alpha x_{t+1}). \quad (32)
\]

In the expression for \( \hat{R}_{t+1} \), we have eliminated \( \hat{y}_t \) using the equilibrium condition (21) and eliminated \( \hat{y}_{t+1} + 1 \) using the definitional relationship

\[
\hat{y}_{t+1} + 1 = \beta^{-1} \exp \left[ -(1 - \alpha)x_{t+1} \right] \hat{z}_{t+1}, \quad (33)
\]

which follows directly from equation (22). To obtain a similar return expression for information set \( J_t \), we define the composite variable \( \overline{z}_{t+1} \equiv \beta \exp \left[ (1 - \alpha) x_{t+1} \right] (\overline{y}_{t+1} + 1) \) and use this definitional relationship and the corresponding equilibrium condition \( \overline{y}_t = E(\overline{z}_{t} | J_t) \) to eliminate \( \overline{y}_t \) and \( \overline{y}_{t+1} + 1 \) from equation (31). In the expression for \( R^*_{t+1} \), we have substituted in \( (y^*_{t+1} + 1) / y^*_t = \beta^{-1} \exp \left[ -(1 - \alpha)x_{t+1} \right] \) from the nonlinear law of motion (18). Notice that

\(^{14}\)Over the period 1871 to 2008, the correlation coefficient between log real price changes and log real equity returns in U.S. annual data is 0.994. LeRoy (1984, p. 186) shows that conditional price variance is numerically very close to the unconditional variance of price changes in a calibrated asset pricing model.
the three return measures differ only by the terms $\hat{z}_{t+1}/E(\hat{z}_{t+1}|H_t)$ and $\bar{z}_{t+1}/E(\bar{z}_{t+1}|J_t)$, which represent the investor’s proportional forecast errors under the different information assumptions. This feature is similar to the excess payoff expressions derived in Section 2 for the simple example involving risk neutrality. Those expressions also differed only in terms of the size of the investor’s forecast errors.\textsuperscript{15}

In Appendix A.2, we show that the approximate law of motion for $\log(\hat{R}_{t+1})$ is
\begin{equation}
\log(\hat{R}_{t+1}) - E[\log(\hat{R}_{t+1})] = \alpha (x_{t+1} - \mu) + a_1 \varepsilon_{t+1}
\end{equation}
where $a_1$ is given by Proposition 1. In Appendix B.2, we show that under $J_t = H_t \cup d_{t+1}$, the approximate law of motion for $\log(\bar{R}_{t+1})$ is
\begin{equation}
\log(\bar{R}_{t+1}) - E[\log(\bar{R}_{t+1})] = n_1 (x_{t+2} - \mu) + (1 - a_1) (x_{t+1} - \mu),
\end{equation}
where $n_1 = a_0 a_1/(1 + a_0)$ is a Taylor-series coefficient with $a_0$ and $a_1$ from Proposition 1. In Appendix C.2, we show that the exact law of motion for $\log(R_{t+1}^*)$ is
\begin{equation}
\log(\hat{R}_{t+1}) - E[\log(\hat{R}_{t+1})] = \alpha (x_{t+1} - \mu).
\end{equation}

Given the above laws of motion for log returns, it is straightforward to compute the following unconditional variances:
\begin{align}
Var[\log(\hat{R}_{t+1})] & = \alpha^2 Var(x_t) + a_1 [a_1 + 2\alpha] \sigma_\varepsilon^2, \quad (37) \\
Var[\log(\bar{R}_{t+1})] & = [(n_1)^2 + (1 - a_1)^2 + 2 n_1 (1 - a_1) \rho] Var(x_t), \quad (38) \\
Var[\log(R_{t+1}^*)] & = \alpha^2 Var(x_t). \quad (39)
\end{align}

### 5.1 Results for Special Cases

LeRoy and Parke (1992) considered the special case of risk neutrality and iid dividend growth. In the present setting, under $\alpha = \rho = 0$, Proposition 1 implies $a_1 = 1$ and equation (14) implies $Var(x_t) = \sigma_\varepsilon^2$. The variance expressions imply the following inequality:
\begin{equation}
Var[\log(R_{t+1}^*)] \leq Var[\log(\hat{R}_{t+1})] \leq Var[\log(\bar{R}_{t+1})], \quad \text{when } \alpha = \rho = 0,
\end{equation}
\footnote{From equation (30), we can see that systematic pessimism would cause the investor’s proportional forecast error to exceed unity on average, thus raising the equilibrium return on equity. Such an effect is explored by Abel (2002).}
where \( a_1 = 1 \) implies \( n_1 = a_0 / (1 + a_0) < 1 \). In this example where \( J_t = H_t \cup d_{t+1} \), the variance of the log return under perfect foresight represents a lower bound of zero while the variance of the log return under information set \( H_t \) represents an upper bound. The results for this special case are analogous to the variance bounds on arithmetic price-changes derived by West (1988b) and Engel (2005) under the assumption of risk neutrality. These authors showed that the variance of arithmetic price-changes declines monotonically as a function of investors’ information about future dividends.

However, it is straightforward to show by counterexample that similar results do not extend to the case where investors are risk averse. Consider the following counterexample when \( \rho = 0 \) but \( \alpha \neq 0 \). We have

\[
\text{Var} \left[ \log(R^*_t) \right] \leq \text{Var} \left[ \log(\hat{R}_{t+1}) \right] \leq \text{Var} \left[ \log(\bar{R}_{t+1}) \right], \quad \text{when } \rho = 0, \tag{41}
\]

where the direction of the second inequality depends on the magnitude of \( \alpha \) and \( n_1 \). Starting from information set \( H_t \) corresponding to \( \log(\hat{R}_{t+1}) \), an increase in investor information can either increase or decrease the log return variance, depending on parameter values.

In the special case of log utility, we have \( \alpha = 1 \) such that \( a_1 = n_1 = 0 \). This case is not a counterexample because it implies

\[
\text{Var} \left[ \log(R^*_t) \right] = \text{Var} \left[ \log(\hat{R}_{t+1}) \right] = \text{Var} \left[ \log(\bar{R}_{t+1}) \right], \quad \text{when } \alpha = 1, \tag{42}
\]

for every specification of \( I_t \). Since the price-dividend ratio is constant under log utility regardless of the information set, return variance is driven solely by the exogenous stochastic process for dividends.

From equations (37) and (39), equality of \( \text{Var}[\log(\hat{R}_{t+1})] \) and \( \text{Var}[\log(R^*_t)] \) can occur not only when \( \alpha = 1 \), but also when \( a_1 + 2\alpha = 0 \). The critical value of \( \alpha \) where \( a_1 + 2\alpha = 0 \) defines a second crossing point at which the size ordering between \( \text{Var}[\log(\hat{R}_{t+1})] \) and \( \text{Var}[\log(R^*_t)] \) again reverses. Consequently, \( \text{Var}[\log(R^*_t)] \) cannot be a lower bound because it may be greater than or less than \( \text{Var}[\log(\hat{R}_{t+1})] \) depending on the value of \( \alpha \).

Solving for the critical value of \( \alpha \) where \( a_1 + 2\alpha = 0 \) can be accomplished analytically using the following approximate expression for the solution coefficient \( a_1 \) in Proposition 1: \( a_1 \simeq (1 - \alpha) / (1 - \rho \beta) \). The approximate expression for \( a_1 \) is derived by assuming \( \exp[(1 - \alpha) \mu + (a_1)^2 \sigma_z^2 / 2] \simeq 1 \) which holds exactly when \( \alpha = 1 \) and remains reasonably accurate for \( \alpha < 10 \). Substituting the approximate expression for \( a_1 \) into the variance equality condition \( a_1 + 2\alpha = 0 \) and then solving for \( \alpha \) yields a second crossing point where \( \text{Var}[\log(\bar{R}_{t+1})] = \text{Var}[\log(R^*_t)] \). The second crossing point is \( \alpha \simeq 1 / (2\rho \beta - 1) \). Positivity of \( \alpha \) requires that the model parameters satisfy \( \rho \beta > 0.5 \).
The intuition for the size ordering reversal is linked to the discounting mechanism. The parameters $\rho$, $\beta$, and $\alpha$ all affect the degree to which future dividend innovations influence the perfect foresight price $p_t^*$ via discounting from the future to the present. When dividends are sufficiently persistent and the investor’s discount factor is sufficiently close to unity such that $\rho \beta > 0.5$, the discounting weights applied to successive future dividend innovations decay more gradually. Since log returns are nearly the same as log price-changes, computation of the log return tends to “difference out” the future dividend innovations, thus shrinking the magnitude of $\text{Var}[\log(R_{t+1}^*)]$ relative to $\text{Var}[\log(\hat{R}_{t+1})]$. In contrast, when $\rho \beta < 0.5$, the discounting weights applied to successive future dividend innovations decay more rapidly, so these terms do not tend to difference out in the log return computation, thus magnifying $\text{Var}[\log(R_{t+1}^*)]$ relative to $\text{Var}[\log(\hat{R}_{t+1})]$.

Given that the size ordering between $\text{Var}[\log(\hat{R}_{t+1})]$ and $\text{Var}[\log(R_{t+1}^*)]$ generally depends on parameter values, these variances cannot represent theoretical bounds for log return volatility. This result should perhaps not be surprising. Unlike the situation with price-dividend ratios, the returns that prevail under information sets $H_t$, $J_t$ and $I_t$ cannot be represented as conditional forecasts of the return that prevails under perfect foresight.$^{16}$

### 5.2 Quantitative Analysis

Figure 2 plots the volatilities of log returns for two different calibrations of the model. In the top panel, we employ the same calibration as Figure 1 with $\rho = -0.1$ to match the autocorrelation of U.S. consumption growth from 1890 to 2008. We see that the volatility of $\log(\hat{R}_{t+1})$ is equal to the volatility of $\log(R_{t+1}^*)$ only when $\alpha = 1$. When agents have no auxiliary information about future dividend realizations (that is, under the information set $H_t$) the model underpredicts return volatility in comparison with the data for any reasonable level of relative risk aversion. The low variance of returns under $H_t$ reflects the specification of near-zero autocorrelation of dividend growth. However, if agents can see dividends one period ahead or an infinite number of periods ahead without error, then the model underpredicts return volatility for relative risk aversion less than 4, but overpredicts it for relative risk aversion greater than 4. Thus if one were using equity return volatility to calibrate the model, and were willing to accept either of these characterizations of investors’ information, then one would conclude that relative risk aversion is about 4. These results are reminiscent of Grossman and Shiller (1981) who employ an informal visual comparison to conclude that a relative risk aversion coefficient of 4 is needed to make the perfect foresight stock price computed from ex post realized dividends look about as volatile as a plot of the S&P 500.

$Lansing (2011)$ shows that a similar variance inequality involving log price changes (rather than log returns) can also be reversed, depending on parameter values. Moreover, he shows that the arithmetic price-change variance bounds derived by West (1988b) and Engel (2005) for the case of risk-neutrality and “cum-dividend” stock prices do not generally extend to the case of ex-dividend prices.
stock price index. Incidentally, it is surprising that the dependence of return volatility on risk aversion that we find is about the same whether investors can see ahead one period or an infinite number of periods ahead.

In the bottom panel of Figure 2, we set \( \rho = 0.7 \) and recalibrate the value of \( \sigma_x \) to maintain the same standard deviation of consumption growth as in the top panel. While this calibration is unrealistic empirically, we consider it to illustrate the point made above that for general parameter values the extreme hypothetical specifications of investors’ information do not establish bounds on return volatility. In this case, the model parameters satisfy \( \rho \beta > 0.5 \) so the return variances based on \( H_t \) and perfect foresight are equal not only when \( \alpha = 1 \), but also when \( \alpha \approx 1/(2\rho \beta - 1) = 2.9 \). When \( \alpha = 2.9 \), we have \( \text{Var}[\log(\tilde{R}_{t+1})] > \text{Var}[\log(\hat{R}_{t+1})] = \text{Var}[\log(R_{t+1}^*]) \), where \( \text{Var}[\log(\tilde{R}_{t+1})] \) is the return variance based on \( J_t = H_t \cup d_{t+1} \). As \( \alpha \) crosses the values 1 and 2.9, the direction of the inequality comparing the volatility of \( \log(\hat{R}_{t+1}) \) to that of \( \log(R_{t+1}^*) \) reverses direction. As observed above, such reversals demonstrate that \( \text{Var}[\log(\tilde{R}_{t+1})] \) and \( \text{Var}[\log(R_{t+1}^*)] \) cannot be bounds for log return volatility.

6 Excess Return Volatility

As a direct counterpart to the concept of excess payoff variance explored in the simple model of section 2, we now consider the implications of the power utility model for the variance of excess returns on equity, i.e., the variance of the equity premium.

In the appendix, we show that the laws of motion for the log risk-free rate under each information set are given by

\[
\log(\hat{R}_{t+1}) - E[\log(\hat{R}_{t+1})] = \alpha \rho (x_t - \mu),
\]

\[
= \alpha \left( x_{t+1} - \mu \right) - \alpha \varepsilon_{t+1},
\] (43)

\[
\log(\tilde{R}_{t+1}) - E[\log(\tilde{R}_{t+1})] = \alpha (x_{t+1} - \mu),
\] (44)

\[
\log(R_{t+1}^*) - E[\log(R_{t+1}^*)] = \alpha (x_{t+1} - \mu).
\] (45)

From the above equations, we see that the risk free rate is identical under information sets \( J_t \) and perfect foresight.\(^{17}\) By definition, the risk-free rate is the return on a one-period bond. Hence, the investor only needs to see consumption/dividends one period ahead in order to see the return on a one-period bond with certainty.

Subtracting the risk-free rate equations from the corresponding equity returns given by equations (34) through (36) yields the following laws of motion for the excess return on equity.

\(^{17}\)From the appendix, we have: \( E[\log(\tilde{R}_{t+1})] = E[\log(\hat{R}_{t+1})] = -\log(\beta) + \alpha \mu \).
under each information set:

\[
\log(\hat{R}_{t+1}) - \log(\hat{R}_{t+1}^f) = (\alpha + a_1) \varepsilon_{t+1} + \frac{1}{2} \left[ \alpha^2 - (a_1)^2 \right] \sigma^2, \tag{46}
\]

\[
\log(\overline{R}_{t+1}) - \log(\overline{R}_{t+1}^f) = [\rho n_1 + 1 - (\alpha + a_1)] (x_{t+1} - \mu) + n_1 \varepsilon_{t+2}, \tag{47}
\]

\[
\log(R_{t+1}^*) - \log(R_{t+1}^{f*}) = 0, \tag{48}
\]

where we have substituted in the expressions for the mean log returns as derived in the appendix.

Equation (48) shows that perfect information about future dividends implies that excess returns are always zero. This is because there is no additional risk to purchasing equity versus a one-period bond when all future dividends are known with certainty. This result is directly analogous to our earlier demonstration in equation (6) that excess payoffs are identically zero under the joint assumptions of perfect foresight and risk neutrality.\(^{18}\)

From equation (48), we have \(V ar \left[ \log\left(\frac{R_{t+1}^*}{R_{t+1}^{f*}}\right) \right] = 0\). Analogous to the simple example of section 2, perfect information about future dividends establishes a theoretical lower bound of zero on excess return volatility even when investors are risk averse. However, equations (46) and (47) imply that information about \(d_{t+1}\) can either increase or decrease excess return variance, depending on the level of risk aversion. Similar to the results for equity returns, the relationship between the variance of excess equity returns and investor information can be non-monotonic. Hence, we cannot establish a theoretical upper bound on excess return variance in the presence of risk aversion.

In the special case when \(\rho = 0\) but \(\alpha \neq 0\), we have

\[
V ar \left[ \log\left(\frac{R_{t+1}^*}{R_{t+1}^{f*}}\right) \right] = 0 \leq V ar \left[ \log\left(\frac{\overline{R}_{t+1}}{\overline{R}_{t+1}^f}\right) \right] \leq V ar [\log(\hat{R}_{t+1}/\hat{R}_{t+1}^f)], \quad \text{when } \rho = 0, \tag{49}
\]

which is similar, but not identical, to the analogous special case (41) derived for return variance. When \(\rho = 0\), we have \(n_1 = a_0 (1 - \alpha) / (1 + a_0)\). It is straightforward to show that \((n_1)^2 < 1\) over the range \(0 < \alpha < 2 + 1/a_0\), whereas \((n_1)^2 > 1\) whenever \(\alpha > 2 + 1/a_0\).

Starting from information set \(H_t\), an increase in investor information can either increase or decrease the variance of excess returns, depending on the value of the risk aversion coefficient \(\alpha\). For lower levels of risk aversion, providing information about \(d_{t+1}\) (moving to information set \(J_t\)) reduces excess return variance such that \(V ar \left[ \log\left(\frac{\overline{R}_{t+1}}{\overline{R}_{t+1}^f}\right) \right] < V ar [\log(\hat{R}_{t+1}/\hat{R}_{t+1}^f)]\).

\(^{18}\)Recall from equation (2) that excess payoffs \(v_{t+1}\) are related to excess returns by the relationship \(v_{t+1}/p_t = R_{t+1} - R_{t+1}^f\). Hence, \(v_{t+1}^* = 0\) also implies \(\log(R_{t+1}^*) - \log(R_{t+1}^{f*}) = 0\).
But for higher levels of risk aversion, providing the same information increases excess return variance such that the variance inequality is reversed.

Figure 3 plots the volatilities of excess log returns on equity for the same two calibrations of the model employed earlier in Figure 2.\textsuperscript{19} The figure confirms that for both calibrations of \( \rho \), the relationship between excess return variance and information about \( d_{t+1} \) is non-monotonic at higher levels of risk aversion.

Veronesi (2000) examines how information about the unobserved drift rate of consumption/dividends affects excess returns on equity in a Lucas-type model with iid dividend growth. In the terminology of our model, he introduces imperfect investor information about the parameter \( \mu \) in a setting with \( \rho = 0 \). He shows (Proposition 4, p. 819) that better information about \( \mu \) can either increase or decrease the variance of excess returns on equity, depending on the level of risk aversion. Interestingly, our results on the ambiguous effect of information about \( d_{t+1} \) on the variance of excess returns shares some of the flavor of Veronesi’s results, even though the concept of information in the two setups is quite different.

7 Mapping to the Campbell-Shiller Framework

Up to this point we have shown that the present-value model with power utility and AR(1) dividend/consumption growth will satisfy the variance bounds for the log price-dividend ratio when the risk aversion coefficient is around 5 or higher. This result contrasts with the findings of excess volatility in the original variance-bounds literature where risk neutrality was routinely assumed. In this section, we examine some other predictions of the power-utility model and show that they differ in important ways from the data.

Campbell and Shiller (1988), Campbell (1991), and Cochrane (1992, 2005) show that a log-linear approximation of the equity return identity (dividend yield plus capital gain) implies that the variance of the log price-dividend ratio must equal the sum of the ratio’s covariances with: (1) future dividend growth rates, (2) future risk-free rates, and (3) future excess returns on equity. This variance decomposition, being derived from an identity rather than a theoretical model, cannot be used to ascertain the theoretical connections between risk aversion, investor information, and stock market volatility. Its use up to now in the finance literature has been to determine the relative empirical importance of the three separate components in explaining the volatility of observed stock prices relative to dividends. However, since the return identity is also valid in theoretical models, it is possible to evaluate our model by performing the variance decomposition analytically and then using the calibrated model to compute the contributions from each of the three components noted above for comparison.

\textsuperscript{19}In computing the volatility of excess log returns in U.S. data, our proxy for the risk-free rate is the one-year real interest rate over the period 1871 to 2008. Data are from Robert Shiller’s website: <www.econ.yale.edu/~shiller/>. 
with the results obtained from U.S. data.

Following the methodology of Campbell and Shiller (1988), the definition of the log equity return under information set $H$, given by equation (30) can be approximated as follows:

$$\log(\hat{R}_{t+1}) \equiv \log (\hat{y}_{t+1} + 1) + x_{t+1} - \log (\hat{y}_t),$$

$$\simeq \kappa_0 + \kappa_1 \log (\hat{y}_{t+1}) + x_{t+1} - \log (\hat{y}_t),$$

where $\kappa_0$ is a constant and $\kappa_1 = \exp \{E \log (\hat{y}_t)\} / \{1 + \exp \{E \log (\hat{y}_t)\}\}$ is a Taylor-series coefficient. Solving equation (50) for $\log (\hat{y}_t)$ and then successively iterating the resulting expression forward to eliminate $\log(\hat{y}_{t+1+j})$ for $j = 0, 1, 2...$ yields the following approximate identity:

$$\log (\hat{y}_t) \simeq \frac{\kappa_0}{1 - \kappa_1} + \sum_{j=0}^{\infty} (\kappa_1)^j \left[ x_{t+1+j} - \log(\hat{R}_{t+1+j}) \right].$$

The return identity shows that movements in the log price-dividend ratio must be accounted for by movements in either future dividend growth rates or future log equity returns. Similar accounting identities can be derived for $\log(\hat{y}_t)$ and $\log(y'_t)$, under information sets $J_t$ and perfect foresight, respectively.

The variables in the approximate identity (51) can be expressed as deviations from their unconditional means while the means are consolidated into the constant term. Multiplying both sides of the resulting expression by $\log (\hat{y}_t) - E [\log (\hat{y}_t)]$ and then taking the unconditional expectation of both sides yields

$$\text{Var} [\log (\hat{y}_t)] = \text{Cov} \left[ \log (\hat{y}_t), \sum_{j=0}^{\infty} (\kappa_1)^j x_{t+1+j} \right] - \text{Cov} \left[ \log (\hat{y}_t), \sum_{j=0}^{\infty} (\kappa_1)^j \log(\hat{R}_{t+1+j}) \right],$$

$$= \text{Cov} \left[ \log (\hat{y}_t), \sum_{j=0}^{\infty} (\kappa_1)^j x_{t+1+j} \right] - \text{Cov} \left[ \log (\hat{y}_t), \sum_{j=0}^{\infty} (\kappa_1)^j \log(\hat{R}_{t+1+j}^f) \right]$$

$$- \text{Cov} \left[ \log (\hat{y}_t), \sum_{j=0}^{\infty} (\kappa_1)^j \log(\hat{R}_{t+1+j}^f / \hat{R}_{t+1+j}^f) \right].$$

(52)

Here, the second version of the expression breaks up the log equity return into two parts: the log risk free rate, denoted by $\log(\hat{R}_{t+1+j}^f)$, and the excess log return on equity, given by $\log(\hat{R}_{t+1+j}^f / \hat{R}_{t+1+j}^f)$. Analogous decompositions can be derived for $\text{Var} [\log (\hat{y}_t)]$ and $\text{Var} [\log (y'_t)]$ which include covariance terms with $\log(\hat{R}_{t+1+j})$ and $\log(R_{t+1+j}^f)$, respectively. The above equation states that the variance of the log price-dividend ratio must be accounted for by the covariance of the log price-dividend ratio with future dividend growth rates, future risk free rates, or future excess returns on equity. The magnitude of each covariance term is a measure of the predictability of future values of dividend growth, risk free rates, or excess returns when the current price-dividend ratio is employed as the sole regressor in a forecasting equation.
For our model, the approximate laws of motion for the log equity return under each information set are given by equations (34) through (36). The corresponding laws of motion for the log risk-free rate are given by equations (43) through (45). Using the approximate laws of motion for the relevant variables, we can analytically compute the three covariance terms in the applicable version of equation (52) for each information set. Details are provided in the appendix. The results of the theoretical variance decomposition are as follows:

\[
\text{Var} \{ \log (b_y_t) \} = a_1 \frac{\rho^2 \text{Var}(x_t)}{1 - \rho \kappa_1} - \frac{a_1 \rho^2 \text{Var}(x_t)}{1 - \rho \kappa_1} - 0, \tag{53}
\]

\[
\text{Var} \{ \log (\bar{y}_t) \} = a_1 \frac{\text{Var}(x_t)}{1 - \rho \kappa_1} - \frac{a_1 \text{Var}(x_t)}{1 - \rho \kappa_1} - \left[ \frac{a_1 (1 - \alpha)}{1 - \rho \kappa_1} - (a_1)^2 \right] \text{Var} (x_t), \tag{54}
\]

\[
\text{Var} \{ \log (y^*_t) \} = \frac{a_1 (1 - \alpha)(1 + \rho \kappa_1 \text{Var}(x_t))}{1 - \rho \kappa_1^2} - \frac{a_1 (1 - \alpha)(1 + \rho \kappa_1 \text{Var}(x_t))}{1 - \rho \kappa_1^2} - 0, \tag{55}
\]

where the three terms in each equation correspond to the three possible sources of variation: (1) future dividend growth rates, (2) future risk free rates, and (3) future excess returns on equity. It should be noted that the expression for the Taylor-series coefficient \( \kappa_1 \) differs slightly across information sets because the unconditional mean of the log price-dividend ratio (the point of approximation for the return identity) depends on the information set.

Equations (53) and (55) show that the variance contribution from excess returns is exactly zero under information sets \( H_t \) and perfect foresight. This result can be understood by examining the laws of motion for excess returns on equity which are given by equations (46) through (48). Equation (46) shows that excess returns are iid under information set \( H_t \), while equation (48) shows that excess returns are identically zero under perfect foresight. In both cases, the covariance between future excess returns and the log price-dividend ratio at time \( t \) is zero. In contrast, equation (47) shows that excess returns under information set \( J_t \) are not iid but instead will inherit the persistence properties of dividend/consumption growth \( x_{t+1} \). When \( \rho = 0 \), we have \( x_{t+1} - \mu = \varepsilon_{t+1} \) and excess returns under information set \( J_t \) will also be iid such that the variance contribution from excess returns will be zero. In this case, the third term in equation (54) will also be zero since \( \rho = 0 \) implies \( a_1 = 1 - \alpha \) from Proposition 1. But even when \( \rho \neq 0 \), the contribution from the third term turns out to be numerically very small for information set \( J_t \).

From equation (46), it is straightforward to see how the inclusion of time-varying risk aversion or time-varying volatility of consumption growth could be used to introduce persistence (or predictibility) into excess returns under information set \( H_t \). Specifically, allowing the risk aversion coefficient \( \alpha \) to depend on lagged values of consumption growth can be interpreted as a reduced-form version of the habit formation model of Campbell and Cochrane (1999). Introducing persistent stochastic variation in the innovation variance \( \sigma^2_\varepsilon \) (which also appears in the expression for \( a_1 \)) would be consistent with the stochastic volatility setup employed by Bansal and Yaron (2004).

22
The theoretical variance decomposition in equations (53) through (55) can be expressed more concisely by dividing both sides of the decomposition by the variance of the log price-dividend ratio for that information set (assuming the variance is non-zero). In this way, the contributions to variance from each source can be expressed as fractions that sum to unity. Details are provided in the appendix. The results of the fractional decomposition are summarized in Table 1.

Table 1: Theoretical variance decomposition for the log price-dividend ratio

<table>
<thead>
<tr>
<th>Information Set</th>
<th>Future Dividend Growth</th>
<th>Future Risk-Free Rates</th>
<th>Future Excess Returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H_t )</td>
<td>( \frac{1}{1-\alpha} )</td>
<td>( -\frac{\alpha}{1-\alpha} )</td>
<td>0</td>
</tr>
<tr>
<td>( J_t )</td>
<td>( \frac{1}{a_1(1-\rho\kappa_1)} )</td>
<td>( -\frac{\alpha}{a_1(1-\rho\kappa_1)} )</td>
<td>( -\left[ \frac{(1-\alpha)}{a_1(1-\rho\kappa_1)} - 1 \right] )</td>
</tr>
<tr>
<td>Perfect Foresight</td>
<td>( \frac{1}{1-\alpha} )</td>
<td>( -\frac{\alpha}{1-\alpha} )</td>
<td>0</td>
</tr>
</tbody>
</table>

Note: For information set \( J_t \), \( a_0 \) and \( a_1 \) are defined by Proposition 1 and \( \kappa_1 = a_0 / (1 + a_0) \).

While the variance of the log price-dividend ratio can differ substantially across the three information sets (as shown previously in Figure 1), the fractions of the variance attributable to each of the three possible sources are exactly the same under information sets \( H_t \) and perfect foresight, and turn out to be only slightly different for information set \( J_t \). Under information sets \( H_t \) and perfect foresight, the variance decomposition depends only on the risk aversion coefficient \( \alpha \). However, for information set \( J_t \), the persistence parameter for dividend growth \( \rho \) also plays a role in the decomposition. As noted above, when \( \rho = 0 \) we have \( a_1 = 1 - \alpha \) from Proposition 1 and the results for information set \( J_t \) are identical to the other two information sets. For our baseline calibration with \( \rho = -0.1 \), the results for information set \( J_t \) are numerically very close to the results for the other two information sets.

Recall that under \( H_t \) and perfect foresight, future excess returns are either iid or zero and hence are not predictable using the log price-dividend ratio at time \( t \). This explains why future excess returns contribute nothing to the variance of the log price-dividend ratio in these two cases. The situation is slightly different for information set \( J_t \). In this case, investors have perfect knowledge of dividends at time \( t+1 \), but they do not have perfect knowledge of equity returns at time \( t+1 \) because they do not know equity prices at \( t+1 \). The investor does have perfect knowledge of the risk-free rate at \( t+1 \) because this depends only on dividend growth from time \( t \) to \( t+1 \), which is known under \( J_t \). The log price-dividend ratio at time \( t \) helps to predict the excess return at \( t+1 \) because it shows the deviation of the ratio from its unconditional mean, which in turn helps to predict the equity price at \( t+1 \).
Under all information sets, the fraction of the variance attributable to future risk free rates is zero when \( \alpha = 0 \), representing risk neutrality. As \( \alpha \) deviates from zero, the agent’s stochastic discount factor becomes more volatile, thus influencing the variance contributions from both future dividend growth rates and future risk-free rates. When \( \alpha = 1 \), the log price-dividend ratio is constant under all information sets and hence there is no variance to be decomposed.

When \( \alpha = 5 \), we have \( 1/(1 - \alpha) = -0.25 \) and \( -\alpha/(1 - \alpha) = 1.25 \). In this case, under \( H_t \) and perfect foresight, \(-25\%\) of the variance of the log price dividend ratio is attributable to movements in future dividend growth, \(125\%\) is attributable to movements in future risk free rates, and \(0\%\) is attributable to movements in future excess returns.\(^{20}\) Under information set \( J_t \) with \( \alpha = 5 \) and \( \rho = -0.1 \), the percentages are the same up to two decimal points. If instead we set \( \rho = 0.7 \), then the contribution from excess returns under \( J_t \) is \(0.38\%\) while the contributions from future dividend growth and future risk free rates are \(-24.91\%\) and \(124.53\%\), respectively. Thus, even a wide variation in the calibrated value of \( \rho \) results in a negligible difference between the variance decomposition under information \( J_t \) and the (common) variance decomposition under \( H_t \) and perfect foresight.

The variance decomposition computed analytically from the power-utility model can be compared to the decomposition obtained from real-world data without imposing any particular model. Cochrane (2005, p. 400) presents an empirical variance decomposition of the log price-dividend ratio for the value-weighted basket of stocks traded on the New York Stock Exchange using annual data for the period 1928 to 1988. The data show that \(-34\%\) of the variance of the log price-dividend ratio is attributable to movements in future dividend growth, while \(138\%\) is attributable to movements in future excess returns (i.e., the sum of future risk free rates and future excess returns). As just noted, with \( \alpha = 5 \) the corresponding percentages from the power-utility model are \(-25\%\) and \(125\%\), respectively, which are not too different from Cochrane’s results. Our model can match the \(-34\%\) figure from Cochrane’s decomposition by setting \( \alpha = 3.9 \), and can match the \(138\%\) figure from Cochrane’s decomposition by setting \( \alpha = 3.6 \). Given the empirical uncertainty surrounding the decomposition in the data, these results are consistent with our earlier finding (plotted in Figure 1) that a risk aversion coefficient of around 5 is needed for the power-utility model to match the volatility of the log price-dividend ratio in U.S. data. So far, so good.

However, if the variance contribution from future equity returns is broken down into separate contributions from future risk-free rates and future excess returns, then the results obtained from the power-utility model are very different from the data. In the data, more than \(100\%\) of the variation in the log price-dividend ratio is attributable to movements in future excess returns, i.e., the third term in equation (52), while almost nothing can be attributed

\(^{20}\)Since the variance decomposition does not require the sources of variation to be orthogonal to one another, the percentage from each source may fall outside the range of \(0\%\) to \(100\%\).
to movements in future risk-free rates. The power-utility model generates the opposite result: more than 100% of the variation in the log price-dividend ratio is attributable to movements in future risk-free rates, while nothing or almost nothing is attributable to movements in future excess returns.

One way in which empirical decomposition manifests itself in the data is the fact that the dividend yield (the inverse of the price-dividend ratio) forecasts excess returns on equity over long horizons, whereas empirical proxies for the risk-free rate do not predict future excess returns, as shown originally by Campbell and Shiller (1988). More recently, Cochrane (2008, p. 1545) obtains a statistically significant long-horizon regression coefficient of 1.23 when forecasting excess returns using the current dividend yield. His result implies that 123% of the dividend yield variance in the data is coming from future excess returns. The power-utility model attributes zero percent (or close to zero percent) of the variance in the dividend yield to future excess returns.

It is important to note that our theoretical variance decomposition for the power-utility model ruled out the presence of bubbles. Specifically, in deriving the approximate return identity (51), we applied the transversality condition

$$\lim_{j \to \infty} (\kappa_1)^j \log \left( y_{t+j} \right) = 0.$$  

In the presence of an explosive rational bubble, the transversality condition would not hold. Recent work by Engsted et al. (2012) shows that the introduction of a periodically-collapsing rational bubble is successful in allowing their theoretical asset pricing model to produce predictability regressions that are very similar to those obtained by Cochrane (2008) for U.S. data. The model employed by Engsted et al. is a special case of the power-utility model considered here with \( \alpha = \rho = 0 \) such that expected returns are constant. Hence, while Cochrane’s empirical results may be interpreted as showing that expected returns in the data are predictable, this interpretation is conditioned on the assumption of no bubbles. If one allows for the presence of bubbles, a version of the power-utility model with constant expected returns can produce data in which the price-dividend ratio has predictive power for future returns.

8 Conclusion

For low levels of investor risk aversion, the volatility of the log price-dividend ratio in U.S. data greatly exceeds the theoretical upper bound implied by the present-value model. Thus we reproduce the result found in the earlier variance-bounds literature. However, for risk aversion coefficients around 5, we find that the volatility of the log price-dividend ratio in long-run U.S. data is reasonably close to the volatility predicted by the power-utility model under the assumption of perfect foresight about future dividends. In other words, the volatility of real-world stock prices relative to dividends is about as one would expect under the assumption that investors can see dividends accurately into the distant future. To be sure, the assumption

---

21 Mehra and Prescott (1985) argue that risk aversion coefficients below 10 are plausible.
that investors have such foresight may be viewed as highly implausible. If one includes in the
null hypothesis the more realistic assumption that investors can foresee dividends at most only
one year into the future, then the volatility of the log price-dividend ratio in the data appears
excessive for any plausible level of risk aversion.

The foregoing conclusion is reminiscent of that of Mehra and Prescott (1985), who found
that they could not explain the observed equity premium in a power-utility model with low
or moderate levels of risk aversion, but could do so with extremely high and implausible
levels of risk aversion—on the order of 50. Similarly, habit formation specifications for utility,
such as that employed by Campbell and Cochrane (1999) to match the equity premium and
other features of the data, imply extremely high coefficients of relative risk aversion. Our
analysis shows that much lower levels of risk aversion will serve the purpose of explaining the
volatility of the log price-dividend ratio in the data if one accepts the notion that investors
can accurately foresee future dividends.

We also demonstrate that previous theoretical bounds on excess payoff variance or arith-
metic price-change variance derived under the assumption of risk neutrality (e.g., West 1988b
and Engel 2005) do not extend to a setting with risk aversion and exponentially growing
dividends, except in some special cases.

Further analysis of the power-utility model raises doubts about whether observed stock
market volatility can be explained with reasonable levels of risk aversion. According to the
power-utility model, the main source of variation in the log price dividend ratio is predictable
variation in future risk-free returns. In contrast, an empirical decomposition finds that the
main source of variation in the log price-dividend ratio is predictable variation in future excess
returns on equity. Thus even though the model can account for the observed volatility of the
price-dividend ratio, it does so by generating an implausibly volatile risk-free rate combined
with an insufficiently predictable excess return on equity. Again, the conclusion is reminiscent
of what other investigations using consumption-based asset pricing theory have found. For
example, Weil (1989) noted that a sufficiently high risk aversion coefficient can succeed in
explaining the equity premium, but at the cost of generating a counterfactually high and
volatile risk-free rate.

Progress on this issue requires a theoretical model that generates variance decompositions
similar to those observed in the data. Allowing for the possibility of rational bubbles is one way
in which the basic power-utility model with low risk aversion might be modified to produce
an empirically-plausible variance decomposition, as demonstrated recently by Engsted et al.
(2012). Alternatively, it would appear that a successful asset pricing model must incorporate
either time-varying risk aversion and/or time-varying volatility of consumption growth. In
this way, excess returns on equity can be made to exhibit significant persistence and volatility,
in contrast to the basic power-utility model which implies that excess returns are iid, as shown
by equation (46).
Campbell and Cochrane (1999) achieve persistent and volatile excess returns on equity by introducing time-varying risk aversion via habit formation. The law of motion for the habit stock in their model is reverse-engineered to deliver a constant risk-free rate, thereby allowing excess returns to make a large contribution to the variance of the log price-dividend ratio, as in the data. However, their calibrated model requires an extremely high coefficient of relative risk aversion to match the various empirical facts—around 80 in the model steady state. Bansal and Yaron (2004) introduce exogenous time-varying volatility in the stochastic processes for consumption and dividend growth which also share a persistent component. In addition, they consider Epstein-Zin preferences which allow the intertemporal elasticity of substitution to be varied independently of the risk aversion coefficient. Nevertheless, their model continues to underpredict the volatility of the log price-dividend ratio in the data even when the risk aversion coefficient is raised to a value of 10.\textsuperscript{22} It therefore remains a challenge for fundamentals-based asset pricing models to explain observed stock market volatility with reasonable levels of risk aversion.

\textsuperscript{22}See Bansal and Yaron (2004), Table IV, p. 1495.
Appendix

A Solution under Information Set \( H_t \)

A.1 Proof of Proposition 1

Iterating ahead the law of motion for \( z_t \) specified in Proposition 1 and taking the conditional expectation implied by the information set \( H_t \) yields

\[
E(z_{t+1}|H_t) = \tilde{y}_t = a_0 \exp \left[ \rho a_1 (x_t - \mu) + \frac{1}{2} (a_1)^2 \sigma^2 \right].
\]  

(A.1)

Substituting the above expression into the first-order condition (22) and then taking logarithms yields

\[
\log (z_t) = F(x_t) = \log (\beta) + (1 - \alpha)x_t + \log \left\{ a_0 \exp \left[ \rho a_1 (x_t - \mu) + \frac{1}{2} (a_1)^2 \sigma^2 \right] + 1 \right\}
\]

\[
\simeq \log (a_0) + a_1 (x_t - \mu),
\]  

(A.2)

where the Taylor-series coefficients \( a_0 \) and \( a_1 \) are given by

\[
\log (a_0) = F(\mu) = \log (\beta) + (1 - \alpha) \mu + \log \left\{ a_0 \exp \left[ \frac{1}{2} (a_1)^2 \sigma^2 \right] + 1 \right\}
\]  

(A.3)

\[
a_1 = \left( \frac{\partial F}{\partial x_t} \right)_{\mu} = 1 - \alpha + \frac{\rho a_1 a_0 \exp \left[ \frac{1}{2} (a_1)^2 \sigma^2 \right]}{a_0 \exp \left[ \frac{1}{2} (a_1)^2 \sigma^2 \right] + 1}.
\]  

(A.4)

Solving equation (A.3) for \( a_0 \) yields

\[
a_0 = \exp \left\{ E[\log (z_t)] \right\} = \frac{\beta \exp \left[ (1 - \alpha) \mu \right]}{1 - \beta \exp \left[ (1 - \alpha) \mu + \frac{1}{2} (a_1)^2 \sigma^2 \right]},
\]  

(A.5)

which can be substituted into equation (A.4) to yield the following nonlinear equation that determines \( a_1 \):

\[
a_1 = 1 - \alpha + \rho a_1 \beta \exp \left[ (1 - \alpha) \mu + \frac{1}{2} (a_1)^2 \sigma^2 \right].
\]  

(A.6)

Rearranging equation (A.6) yields the expression shown in Proposition 1. There are two solutions, but only one solution satisfies the condition \( \beta \exp \left[ (1 - \alpha) \mu + \frac{1}{2} (a_1)^2 \sigma^2 \right] < 1 \), which is verified after solving (A.6) using a nonlinear equation solver. ■
A.2 Asset Pricing Moments

This section briefly outlines the derivation of equations (24) and (37). Taking the unconditional expectation of \( \log (\tilde{y}_t) \) in equation (23) yields

\[
E \left[ \log (\tilde{y}_t) \right] = \log (a_0) + \frac{1}{2} (a_1)^2 \sigma^2_{\tilde{\varepsilon}}. \tag{A.7}
\]

We then have

\[
\log (\tilde{y}_t) - E [\log (\tilde{y}_t)] = a_1 \rho (x_t - \mu), \tag{A.8}
\]

which in turn implies

\[
Var [\log (\tilde{y}_t)] = (a_1 \rho)^2 Var (x_t). \tag{A.9}
\]

As described in the text, the equity return (30) implied by the information set \( H_t \) can be rewritten as

\[
\hat{R}_{t+1} = \beta^{-1} \exp (\alpha \cdot x_{t+1}) \left[ \frac{\hat{z}_{t+1}}{E(z_{t+1}|H_t)} \right]. \tag{A.10}
\]

Substituting in \( E(z_{t+1}|H_t) \) from equation (A.1) and \( \hat{z}_{t+1} = a_0 \exp [a_1 (x_{t+1} - \mu)] \) from Proposition 1 and then taking the unconditional mean of \( \log(\hat{R}_{t+1}) \) yields

\[
E[\log(\hat{R}_{t+1})] = -\log (\beta) + \alpha \mu - \frac{1}{2} (a_1)^2 \sigma^2_{\tilde{\varepsilon}}. \tag{A.11}
\]

We then have

\[
\log(\hat{R}_{t+1}) - E[\log(\hat{R}_{t+1})] = \alpha (x_{t+1} - \mu) + a_1 \varepsilon_{t+1}, \tag{A.12}
\]

which in turns implies

\[
Var[\log(\hat{R}_{t+1})] = \alpha^2 Var (x_t) + (a_1)^2 \sigma^2_{\tilde{\varepsilon}} + 2 \alpha a_1 Cov (x_{t+1}, \varepsilon_{t+1}) \bigg|_{= \sigma^2_{\tilde{\varepsilon}}} = \sigma^2_{\tilde{\varepsilon}}. \tag{A.13}
\]

The log risk free rate is determined by the following first-order condition

\[
\log(\hat{R}_{t+1}^{f}) = -\log \left\{ E \left[ \beta \left( \frac{c_{t+1}}{c_t} \right)^{-\alpha} |H_t \right] \right\},
\]

\[
= -\log \left\{ E \left[ \beta \exp (-\alpha \cdot x_{t+1}) |H_t \right] \right\},
\]

\[
= -\log (\beta) + \alpha [px_t + (1 - \rho) \mu] - \frac{1}{2} \alpha^2 \sigma^2_{\tilde{\varepsilon}}, \tag{A.14}
\]

where we have made the substitution \( c_{t+1}/c_t = d_{t+1}/d_t = \exp (x_{t+1}) \) and then inserted the law of motion for \( x_{t+1} \) from equation (12) before taking the conditional expectation. Taking the unconditional mean of \( \log(\hat{R}_{t+1}^{f}) \) and then subtracting the unconditional mean from equation (A.14) yields the law of motion (43).
A.3 Variance Decomposition

For information set \( H_t \), we have

\[
\kappa_1 = \frac{\exp \left[ E \log (\tilde{y}_t) \right]}{1 + \exp \left[ E \log (\tilde{y}_t) \right]} = \frac{a_0 \exp \left[ \frac{1}{2} (a_1)^2 \sigma_{\tilde{y}}^2 \right]}{1 + a_0 \exp \left[ \frac{1}{2} (a_1)^2 \sigma_{\tilde{y}}^2 \right]} = \beta \exp \left[ (1 - \alpha) \mu + \frac{1}{2} (a_1)^2 \sigma_{\tilde{y}}^2 \right], \quad (A.15)
\]

where we have made use of equations (A.5) and (A.7).

Using the law of motion for \( \log (b_{yt}) \) given by (A.8) and the law of motion for \( x_{t+1} \) given by (12), we can compute the following covariance which is the first term in equation (52):

\[
\text{Cov} \left[ \log (\tilde{y}_t), \sum_{j=0}^{\infty} (\kappa_1)^j x_{t+1+j} \right] = E \left\{ a_1 \rho(x_t - \mu) (x_{t+1} - \mu) + \kappa_1 a_1 \rho(x_t - \mu) (x_{t+2} - \mu) + (\kappa_1)^2 a_1 \rho (x_t - \mu) (x_{t+3} - \mu) + \ldots \right\},
\]

\[
= a_1 \rho^2 \text{Var} (x_t) \left\{ 1 + \rho \kappa_1 + (\rho \kappa_1)^2 + (\rho \kappa_1)^3 + \ldots \right\},
\]

\[
= \frac{a_1 \rho^2 \text{Var} (x_t)}{1 - \rho \kappa_1}. \quad (A.16)
\]

Similarly, using the law of motion for \( \log (\tilde{R}_{t+1}^f) \) given by (43) we can compute the following covariance which is the second term in equation (52):

\[
-\text{Cov} \left[ \log (\tilde{y}_t), \sum_{j=0}^{\infty} (\kappa_1)^j \log (\tilde{R}_{t+1}^f) \right] = -E \left\{ a_1 \rho (x_t - \mu) \alpha \rho (x_t - \mu) + \kappa_1 a_1 \rho (x_t - \mu) \alpha \rho (x_{t+2} - \mu) + (\kappa_1)^2 a_1 \rho (x_t - \mu) \alpha \rho (x_{t+3} - \mu) + \ldots \right\},
\]

\[
= -\alpha a_1 \rho^2 \text{Var} (x_t) \left\{ 1 + \rho \kappa_1 + (\rho \kappa_1)^2 + (\rho \kappa_1)^3 + \ldots \right\},
\]

\[
= -\frac{\alpha a_1 \rho^2 \text{Var} (x_t)}{1 - \rho \kappa_1}. \quad (A.17)
\]

The law of motion for excess returns (46) shows that excess returns are \( iid \) under information set \( H_t \). Hence, the third covariance term in equation (52) is identically zero. Dividing both sides of equation (52) by \( \text{Var} [\log (\tilde{y}_t)] \) from (A.9) and then substituting in the appropriate moments yields

\[
1 = \frac{1}{a_1 (1 - \rho \kappa_1)} - \frac{\alpha}{a_1 (1 - \rho \kappa_1)} - 0,
\]

\[
= \frac{1}{1 - \alpha} - \frac{\alpha}{1 - \alpha} - 0, \quad (A.18)
\]

where we make use of \( (1 - \rho \kappa_1) = (1 - \alpha)/a_1 \) from equations (A.6) and (A.15).
B  Solution under Information Set  \( J_t = H_t \cup d_{t+1} \)

B.1  Characterizing  \( \bar{y}_t \)

Iterating ahead the first-order condition (9) and then imposing the equilibrium relationship  \( c_t = d_t \) for all \( t \) yields

\[
\bar{p}_t = \beta \left( \frac{d_{t+1}}{d_t} \right)^{-\alpha} (d_{t+1} + \hat{p}_{t+1}) , \tag{B.1}
\]

where  \( \hat{p}_{t+1} \) is the equilibrium price conditional on the information set  \( H_{t+1} \).

Dividing both sides of equation (B.1) by  \( d_t \) yields the following expression for  \( \bar{y}_t \equiv \bar{p}_t / d_t \):

\[
\bar{y}_t = \beta \exp ((1 - \alpha) x_{t+1}) (1 + \hat{y}_{t+1}) ,
\]

\[
= \hat{z}_{t+1} , \tag{B.2}
\]

where the second equality follows directly from the definition (20).

Given that  \( \bar{y}_t = \hat{z}_{t+1} \) from equation (B.2) and  \( \hat{y}_t = E(\hat{z}_{t+1} | H_t) \) from equation (A.1), we then have  \( \hat{y}_t = E(\bar{y}_t | H_t) \) which implies  \( Var(\hat{y}_t) \leq Var(\bar{y}_t) \).

B.2  Asset Pricing Moments

This section outlines the derivation of equations (27) and (38). From equations (A.2) and (B.2) we have the following approximate law of motion for  \( \bar{y}_t \)

\[
\bar{y}_t = \hat{z}_{t+1} \simeq a_0 \exp [a_1 (x_{t+1} - \mu)] , \tag{B.3}
\]

which implies  \( E[\log(\bar{y}_t)] = \log(a_0) < E[\log(\hat{y}_t)] \). The above expression implies

\[
Var[\log(\bar{y}_t)] = (a_1)^2 Var(x_t) . \tag{B.4}
\]

The equity return (31) implied by the information set  \( J_t \) can be rewritten as

\[
\bar{R}_{t+1} = \exp(x_{t+1}) \left[ \frac{\hat{z}_{t+2} + 1}{\hat{z}_{t+1}} \right] , \tag{B.5}
\]

where we have eliminated both  \( \bar{y}_t \) and  \( \bar{y}_{t+1} \) using equation (B.2). The approximate law of motion for  \( \hat{z}_{t+1} \) is given by equation (B.3). An approximate law of motion for  \( \hat{z}_{t+2} + 1 \) is given by

\[
\hat{z}_{t+2} + 1 \simeq n_0 \exp [n_1 (x_{t+2} - \mu)] , \tag{B.6}
\]

where  \( n_0 = 1 + a_0 \) and  \( n_1 = a_0 a_1 / (1 + a_0) \) are Taylor-series coefficients.
Substituting equations (B.3) and (B.6) into (B.5) and then taking the unconditional mean of \( \log(\bar{R}_{t+1}) \) yields

\[
E \left[ \log(\bar{R}_{t+1}) \right] = \log(n_0/a_0) + \mu,
\]

\[
= -\log(\beta) + \alpha \mu \tag{B.7}
\]

We then have

\[
\log(\bar{R}_{t+1}) - E \left[ \log(\bar{R}_{t+1}) \right] = n_1 (x_{t+2} - \mu) + (1 - a_1) (x_{t+1} - \mu). \tag{B.8}
\]

Squaring both sides of equation (B.8) and taking the unconditional mean yields the expression for \( \text{Var} \left[ \log(\bar{R}_{t+1}) \right] \) shown in equation (38).

The log risk free rate is determined by the following first-order condition

\[
\log(\bar{R}_{t+1}^f) = -\log \left\{ E \left[ \beta \left( \frac{c_{t+1}}{c_t} \right)^{-\alpha} | J_t \right] \right\},
\]

\[
= -\log \left\{ \exp \left( -\alpha x_{t+1} \right) \right\},
\]

\[
= -\log(\beta) + \alpha x_{t+1}, \tag{B.9}
\]

where we have made the substitution \( c_{t+1}/c_t = d_{t+1}/d_t = \exp(x_{t+1}) \). Given that \( J_t = H_t \cup d_{t+1} \), the investor has perfect knowledge of \( x_{t+1} \) at time \( t \) so we may drop the conditional expectation. Taking the unconditional mean of \( \log(\bar{R}_{t+1}^f) \) and then subtracting the unconditional mean from equation (B.9) yields the law of motion (44).

### B.3 Variance Decomposition

For information set \( J_t \), we have

\[
\kappa_1 = \frac{\exp \left[ E \log(\bar{y}_t) \right]}{1 + \exp \left[ E \log(\bar{y}_t) \right]} = \frac{a_0}{1 + a_0} = \frac{\beta \exp[(1-\alpha)\mu]}{1 + \beta \exp[(1-\alpha)\mu] - \beta \exp[(1-\alpha)\mu + \frac{1}{2}(a_1)^2 \sigma^2]}, \tag{B.10}
\]

where we have made use of equations (A.5) and (B.3). Note that \( \kappa_1 \) here differs from that under information set \( H_t \), as shown in equation (A.15).

Using the law of motion for \( \log(\bar{y}_t) \) given by (B.3) and the law of motion for \( x_{t+1} \) given by (12), we can compute the following covariance which is the first term in the variance
decomposition:

\[
\text{Cov} \left[ \log (\bar{y}_t), \sum_{j=0}^{\infty} (\kappa_1)^j x_{t+1+j} \right] = E \left\{ a_1 (x_{t+1} - \mu) (x_{t+1} - \mu) + \kappa_1 a_1 (x_{t+1} - \mu) (x_{t+2} - \mu) \right.
\]
\[+ (\kappa_1)^2 a_1 (x_{t+1} - \mu) (x_{t+3} - \mu) + \ldots \right\},
\]
\[= a_1 \text{Var} (x_t) \left\{ 1 + \rho \kappa_1 + (\rho \kappa_1)^2 + (\rho \kappa_1)^3 + \ldots \right\},
\]
\[= \frac{a_1 \text{Var} (x_t)}{1 - \rho \kappa_1}. \quad (B.11)
\]

Similarly, using the law of motion for \(\log(R_{t+1})\) given by (44) we can compute the following covariance which is the second term in the variance decomposition:

\[
-Cov \left[ \log (\bar{y}_t), \sum_{j=0}^{\infty} (\kappa_1)^j \log (\bar{R}_{t+1}^f) \right] = -E \left\{ a_1 (x_{t+1} - \mu) \alpha (x_{t+1} - \mu)
\right.
\]
\[+ \kappa_1 a_1 (x_{t+1} - \mu) \alpha (x_{t+2} - \mu)
\]
\[+ (\kappa_1)^2 a_1 (x_{t+1} - \mu) \alpha (x_{t+3} - \mu) + \ldots \right\},
\]
\[= \alpha a_1 \text{Var} (x_t) \left\{ 1 + \rho \kappa_1 + (\rho \kappa_1)^2 + (\rho \kappa_1)^3 + \ldots \right\},
\]
\[= \frac{\alpha a_1 \text{Var} (x_t)}{1 - \rho \kappa_1}. \quad (B.12)
\]

Using the law of motion for excess returns (47), the third term in the variance decomposition is given by

\[
-Cov \left[ \log (\bar{y}_t), \sum_{j=0}^{\infty} (\kappa_1)^j \log \left( \frac{\bar{R}_{t+1+j}}{\bar{R}_{t+1}} \right) \right] = -E \left\{ a_1 (x_{t+1} - \mu) (1 - \alpha - \alpha_1 + \rho \kappa_1) (x_{t+1} - \mu)
\right.
\]
\[+ \kappa_1 a_1 (x_{t+1} - \mu) (1 - \alpha - \alpha_1 + \rho \kappa_1) (x_{t+2} - \mu) + \ldots \right\},
\]
\[= a_1 (1 - \alpha - \alpha_1 + \rho \kappa_1) \text{Var} (x_t) \left\{ 1 + \rho \kappa_1 + (\rho \kappa_1)^2 + \ldots \right\},
\]
\[= \frac{a_1 (1 - \alpha - \alpha_1 + \rho \kappa_1) \text{Var} (x_t)}{1 - \rho \kappa_1}
\]
\[= \frac{a_1 [1 - \alpha - \alpha_1 (1 - \rho \kappa_1)] \text{Var} (x_t)}{1 - \rho \kappa_1}. \quad (B.13)
\]

where we note that \(n_1 = \kappa_1 a_1\).
Dividing both sides of the variance decomposition by $\text{Var} [\log (y_t)]$ from equation (B.4) and then substituting in the appropriate moments yields

$$1 = \frac{1}{a_1 (1 - \rho \kappa_1)} - \frac{\alpha}{a_1 (1 - \rho \kappa_1)} - \left[ \frac{(1 - \alpha)}{a_1 (1 - \rho \kappa_1)} - 1 \right],$$  \hspace{1cm} (B.14)$$

where equations (A.6) and (B.10) imply $(1 - \rho \kappa_1) \simeq (1 - \alpha) / a_1$ only when $\rho \simeq 0$.

\section*{C Solution under Perfect Foresight}

\subsection*{C.1 Log-linearized Law of Motion}

Taking logarithms of the nonlinear law of motion (18) yields

$$\log (y^*_t) = G \left[ x_{t+1}, \log (y^*_{t+1}) \right] = \log (\beta) + (1 - \alpha) x_{t+1} + \log \left\{ \exp \left[ \log (y^*_{t+1}) \right] + 1 \right\}$$

$$\simeq \log (b_0) + b_1 (x_{t+1} - \mu) + b_2 \left[ \log (y^*_{t+1}) - \log (b_0) \right],$$  \hspace{1cm} (C.1)$$

where the Taylor-series coefficients $b_0$, $b_1$, and $b_2$ are given by

$$\log (b_0) = G [\mu, \log (b_0)] = \log (\beta) + (1 - \alpha) \mu + \log [b_0 + 1],$$  \hspace{1cm} (C.2)$$

$$b_1 = \frac{\partial G}{\partial x_t} \bigg|_{\mu, \log (b_0)} = 1 - \alpha,$$  \hspace{1cm} (C.3)$$

$$b_2 = \frac{\partial G}{\partial \log (y^*_{t+1})} \bigg|_{\mu, \log (b_0)} = \frac{b_0}{b_0 + 1}.$$  \hspace{1cm} (C.4)$$

Solving equation (C.2) for the unconditional mean $b_0$ yields

$$b_0 = \exp \left\{ E \left[ \log (y^*_t) \right] \right\} = \frac{\beta \exp [(1 - \alpha) \mu]}{1 - \beta \exp [(1 - \alpha) \mu]},$$  \hspace{1cm} (C.5)$$

which can be substituted into equation (C.4) to obtain the following expression for $b_2$:

$$b_2 = \beta \exp [(1 - \alpha) \mu].$$  \hspace{1cm} (C.6)$$

Subtracting $\log (b_0) = E \left[ \log (y^*_t) \right]$ from both sides of the approximate law of motion (C.1) and then substituting for $b_1$ and $b_2$ from (C.3) and (C.6) yields equation (28).
C.2 Asset Pricing Moments

This section outlines the derivation of equations (29) and (39). Squaring both sides of equation (28) and then taking the unconditional mean to obtain the variance yields

\[
\text{Var} [\log (y_t^*)] = \frac{(1 - \alpha)^2 \text{Var} (x_t) + 2(1 - \alpha) \beta \exp [(1 - \alpha) \mu] \text{Cov} [\log (y_t^*), x_t]}{1 - \beta^2 \exp [2(1 - \alpha) \mu]}.
\]  
(C.7)

The next step is to compute \( \text{Cov} [\log (y_t^*), x_t] \) which appears in equation (C.7). Starting from equation (28), we have

\[
\text{Cov} [\log (y_t^*), x_t] = (1 - \alpha) \text{Cov} (x_{t+1}, x_t) + \beta \exp [(1 - \alpha) \mu] \text{Cov} [\log (y_t^*), x_t],
\]  
(C.8)

\[
\text{Cov} [\log (y_{t+1}^*), x_t] = (1 - \alpha) \text{Cov} (x_{t+1}, x_{t-1}) + \beta \exp [(1 - \alpha) \mu] \text{Cov} [\log (y_{t+2}^*), x_t],
\]  
(C.9)

and so on for \( \text{Cov} [\log (y_{t+j}^*), x_t], j = 1, 2, 3, \ldots \) By repeated substitution to eliminate \( \text{Cov} [\log (y_{t+j}^*), x_t] \) and then applying a transversality condition, we obtain the following expression:

\[
\text{Cov} [\log (y_t^*), x_t] = (1 - \alpha) \text{Cov} (x_t, x_{t-1}) \sum_{j=0}^{\infty} \{\rho \beta \exp [(1 - \alpha) \mu]\}^j = \frac{(1 - \alpha) \text{Cov} (x_t, x_{t-1})}{1 - \rho \beta \exp [(1 - \alpha) \mu]} = \frac{(1 - \alpha) \rho \text{Var} (x)}{1 - \rho \beta \exp [(1 - \alpha) \mu]},
\]  
(C.10)

where the infinite sum converges provided that \( \rho \beta \exp [(1 - \alpha) \mu] < 1 \). Substituting equation (C.10) into equation (C.7), then simplifying yields equation (29).

The perfect foresight return (32) can be rewritten as

\[
R_{t+1}^* = \beta^{-1} \exp (\alpha x_{t+1}),
\]  
(C.11)

where we have substituted in \((y_{t+1}^* + 1)/y_t^* = \beta^{-1} \exp [-(1 - \alpha) x_{t+1}]\) from the exact nonlinear law of motion (18). Taking the unconditional expectation of \( \log (R_{t+1}^*) \) yields

\[
E [\log (R_{t+1}^*)] = -\log (\beta) + \alpha \mu.
\]  
(C.12)

We then have

\[
\log (R_{t+1}^*) - E [\log (R_{t+1}^*)] = \alpha (x_{t+1} - \mu),
\]  
(C.13)
which in turns implies the unconditional variance (39).

The log risk free rate is determined by the following perfect-foresight version of the first-order condition

\[
\log(R_{t+1}^f) = - \log \left\{ \beta \left( \frac{c_{t+1}}{c_t} \right)^{-\alpha} \right\},
\]

\[
= - \log \{ \beta \exp (-\alpha x_{t+1}) \},
\]

\[
= - \log (\beta) + \alpha x_{t+1}, \tag{C.14}
\]

where we have made the substitution \(c_{t+1}/c_t = d_{t+1}/d_t = \exp (x_{t+1})\). Taking the unconditional mean of \(\log(R_{t+1}^f)\) and then subtracting the unconditional mean from equation (C.14) yields the law of motion (45).

C.3 Variance Decomposition

Under perfect foresight, we have

\[
\kappa_1 = \frac{\exp \{ E [\log (y_t^*)]\}}{1 + \exp \{ E [\log (y_t^*)]\}} = \frac{b_0}{1 + b_0} = \beta \exp [(1 - \alpha) \mu], \tag{C.15}
\]

where we have made use of equation (C.5).

Using the law of motion for \(\log (y_t^*)\) given by (C.1) and the law of motion for \(x_{t+1}\) given by (12), we can compute the following covariance which is the first term in the variance decomposition:

\[
Cov \left[ \log (y_t^*), \sum_{j=0}^{\infty} (\kappa_1)^j x_{t+1+j} \right] = E \left\{ (1 - \alpha) (x_{t+1} - \mu) (x_{t+1} - \mu) + b_2 [\log (y_{t+1}^*/b_0)] (x_{t+1} - \mu) + \kappa_1 (1 - \alpha) (x_{t+1} - \mu) (x_{t+2} - \mu) + \kappa_1 b_2 \log (y_{t+1}^*/b_0) (x_{t+2} - \mu) + \kappa_1^2 (1 - \alpha) (x_{t+1} - \mu) (x_{t+3} - \mu) + \kappa_1^2 b_2 \log (y_{t+1}^*/b_0) (x_{t+3} - \mu) + \ldots \right\}
\]

\[
= (1 - \alpha) Var (x_t) \left\{ 1 + \rho \kappa_1 + (\rho \kappa_1)^2 + \ldots \right\} + \kappa_1^2 (1 - \alpha) Var (x_t) \left\{ 1 + \rho \kappa_1 + (\rho \kappa_1)^2 + \ldots \right\} + \kappa_1 \mu \log (y_t^*) \right\} + \kappa_1 \mu x_t \right\} \left\{ 1 + (\kappa_1)^2 + (\kappa_1)^4 + \ldots \right\}
\]

\[
\leq \frac{(1 - \alpha) Var (x_t)}{1 - \rho \kappa_1} \left\{ 1 + (\kappa_1)^2 + (\kappa_1)^4 + \ldots \right\} + \kappa_1 Cov \left[ \log (y_t^*), x_t \right] \left\{ 1 - (\kappa_1)^2 + (\kappa_1)^4 + \ldots \right\}
\]

\[
= \frac{(1 - \alpha) (1 + \rho \kappa_1) Var (x_t)}{[1 - (\kappa_1)^2] (1 - \rho \kappa_1)}, \tag{C.16}
\]
where we have substituted in $\text{Cov}[\log(y_t^*), x_t]$ from equation (C.10).

Following the same methodology and using the law of motion for $\log(R_{t+1}^*)$ given by (45), we can compute the following covariance which is the second term in the variance decomposition:

$$-\text{Cov} \left[ \log(\tilde{y}_t), \sum_{j=0}^{\infty} (\kappa_1)^j \log(R_{t+1+j}^*) \right] = -\frac{\alpha (1 - \alpha) (1 + \rho \kappa_1) \text{Var}(x_t)}{[1 - (\kappa_1)^2] (1 - \rho \kappa_1)} \quad (C.17)$$

The law of motion for excess returns (48) shows that excess returns are identically zero under perfect foresight. Hence, the third term in the variance decomposition is identically zero. Dividing both sides of the variance decomposition by $\text{Var}[\log(y_t^*)]$ from equation (29) and then substituting in the appropriate moments yields

$$1 = \frac{1}{1 - \alpha} - \frac{\alpha}{1 - \alpha} - 0, \quad (C.18)$$

where we make use of $\rho \beta \exp[(1 - \alpha)\mu] = \rho \kappa_1$ and $\beta^2 \exp[2(1 - \alpha)\mu] = (\kappa_1)^2$ from equation (C.15).
References


Figure 1: The log price-dividend ratio in U.S. data exhibits excess volatility for $\alpha < 5$. 

![Graph showing volatility of log price-dividend ratio](image-url)
Figure 2: The present-value model does not impose bounds on equity return volatility in general settings involving risk aversion.
Figure 3: Perfect information about future dividends imposes a lower bound of zero on excess return volatility. However, providing some information about future dividends can either increase or decrease excess return volatility, depending on the level of risk aversion.