Risk Aversion and Stock Price Volatility

Kevin J. Lansing*
Federal Reserve Bank of San Francisco

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

May 25, 2011

Abstract

This paper employs a standard asset pricing framework with power utility to derive model-predicted volatility measures for the price-dividend ratio in a setting that allows for varying degrees of investor information about future dividends. When comparing the model predictions to the data, we find evidence of excess volatility in long-run U.S. stock price data for relative risk aversion coefficients below 5. For higher degrees of risk aversion, the evidence for excess volatility is less clear. We also examine the degree to which movements in the model 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. We show that the theoretical variance decomposition differs in important ways from the data. Thus even though the model can account for the observed volatility of the price-dividend ratio in the data, it does so by generating an implausibly volatile risk-free rate combined with an insufficiently forecastable excess return on equity.

Keywords: Asset Pricing, Excess Volatility, Variance Bounds, Risk Aversion, Present-Value Model.

JEL Classification: E44, G12.

* 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
Risk Aversion and Stock Price Volatility

1 Introduction

The variance-bounds tests of stock price volatility reported in the mid-1970s and later raised doubts about whether stock prices could be accurately represented as the present value of expected future dividends with a constant discount factor. Volatility in the data appeared to exceed the levels implied by the discounted-dividend model with a constant discount factor. How were these findings to be reconciled with the conclusions of the earlier efficient markets literature supporting the discounted-dividend model? A number of econometric problems with the variance-bounds tests were unearthed, but it turned out that correcting these problems did not eliminate the appearance of excess volatility.\(^1\)

The debate ended—somewhat abruptly—in the 1990s when economists using analytical methods unrelated to those employed in the variance-bounds tests began reaching conclusions that reversed those of the earlier efficient-markets literature. Campbell and Shiller (1988), for example, employed a log-linear decomposition of the equity return identity that turned out to be extremely fruitful. Using this decomposition, they showed that observed stock price volatility can be attributed mainly to the fact that future returns contain a predictable component, contrary to the implications of the constant discount factor model employed in the variance-bounds literature. Predictable future returns imply stochastic discount rates, indicative of risk aversion, whereas the constant discount rate model reflected the assumption of risk neutrality. Hence, a finding of apparent excess volatility could be induced by the misspecification of constant discount rates.

This paper, like that of Campbell and Shiller (1988), makes use of log-linear methods to examine the connection between risk aversion and stock price 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.

The fact that two distinct analytical procedures reach similar conclusions, i.e., rejection of the constant discount-rate version of the present-value model, leaves open the question of which method is to be preferred for research on asset pricing. Some take the view that the Campbell-Shiller method, which is not tied to a particular asset pricing model, is less prone to misspecification. Perhaps so, but the fact that analysts using variance-bounds tests came to the common conclusion much earlier than others suggests that the variance-bounds tests

\(^1\)For summaries of this literature, see West (1988a), Gilles and LeRoy (1991), Shiller (2003), and LeRoy (2010).
focus attention on essential considerations that might otherwise be overlooked.

Let us see what these considerations might be. 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 (i.e., marginal) process for dividends taken as given, but 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 of future dividends. Existence of such auxiliary information variables has implications for price volatility. These implications are ignored in many asset pricing models, but they play a major role in the variance-bounds tests.

In the risk neutral case, an increase in the amount of investors’ auxiliary information about future dividend innovations will raise the unconditional variance of stock prices but lower the variance of excess payoffs or price changes.2 The upper bound on unconditional price variance, corresponding to a lower bound of zero on payoff variance, is reached when investors are assumed to have perfect foresight about the entire future path of dividends. In contrast, the upper bound on payoff variance is reached in the opposite case when investors have no auxiliary information about future dividend innovations, i.e., when investors only know current and past dividends. 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.

The point is that a comparison of actual and perfect foresight stock prices, as is inherent in the variance-bounds tests, forces the analyst to focus on the degree to which investors possess information over and above that contained in current and past dividends. The variance bounds are generated by imposing extreme hypothetical specifications for investors’ information. Given the explicit focus on information assumptions, variance-bounds tests remain a useful analytical tool for assessing the success or failure of the present-value model. Of course, we do not suggest restricting attention to these methods.

When economists talk about stock price volatility it is not clear whether unconditional or conditional variance best corresponds to what they have in mind.3 The frequent use of terms like “choppiness” or “smoothness” in describing 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 just noted that the unconditional and conditional variances are affected in opposite directions by variables measuring the extent of investors’ auxiliary information

---

2This basic result is established in LeRoy and Porter (1981), West (1988b), and LeRoy (1996).
3LeRoy (1984) and Kleidon (1986) use numerical examples to illustrate the idea that different conclusions may be drawn from considering conditional variances rather than unconditional variances.
implies that this lack of clarity is an important problem. We do not take the position here that either of these variance measures is superior to the other; instead we will determine the implications of the present-value model, or lack thereof, for both.

A major shortcoming of the early variance-bounds tests was the maintained assumption that equity returns follow a white-noise process. While this assumption may be justified when investors are risk neutral, the time-series behavior of equity returns will be more complex than a white-noise process if investors are risk averse. The assumption of risk neutrality in the early variance-bounds tests appeared acceptable at the time because the empirical literature on stock market efficiency had been based on exactly the same assumption, although this was seldom made explicit. However, the development of quantitative consumption-based asset pricing models, as in Mehra and Prescott (1985), led analysts to a heightened appreciation of the potential importance of risk aversion.

Risk aversion was recognized early as a possible cause for findings of apparent excess volatility in stock prices (see Grossman and Shiller 1981 and LeRoy and LaCivita 1981); at least in some settings risk aversion increases the implied volatility of stock prices relative to the risk-neutral case. This is so because risk-averse investors wish to smooth consumption over time. They do this by trying to acquire financial assets in good times and liquidate them otherwise. In such settings, market clearing implies that stock prices will be higher in good times and lower in bad times by a factor that depends on risk aversion, relative to the risk-neutral case.

These arguments, however, were incomplete. Establishing that risk aversion may affect actual stock price 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 implies high price volatility in some settings, it may or may not imply excess price volatility. It remained unclear whether the existence of risk aversion could alter the proposition that the variance computed from perfect foresight prices represents an upper bound for the variance computed from actual prices.

This paper compares the price and return volatilities computed from actual 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, we find evidence of excess volatility in long-run U.S. stock price 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. To the extent that this assumption is viewed as implausible, it follows that price volatility is excessive in that case as well.

---

4See LeRoy (1973) and Lucas (1978).
In log-linear settings, as here, equity return variance is the analog to the concept of excess payoff variance or price-change variance examined by LeRoy and Porter (1981), West (1988b), and LeRoy (1996). We show by counterexample that when investors are risk averse, the return variance analogs to the earlier LeRoy-Porter-West results may not apply; equity return variance is not necessarily a monotone decreasing function of investors’ information. Therefore, the extreme hypothetical specifications of investors’ information do not necessarily provide bounds on equity return variance. We demonstrate that, despite the absence of theoretical variance bounds for equity returns, the present-value model can 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.

Last, we derive a theoretical variance decomposition for the model price-dividend ratio under different information sets. Specifically, 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 show that the coefficient of relative risk aversion is the key parameter that governs the fraction of the price-dividend ratio’s variance that is attributable to either future dividend growth rates or future risk-free rates. 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.

The variance contribution from future excess returns in the power utility model turns out to be either 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 with the model’s predictions. Recent contributions to the theoretical literature that go beyond the power utility model have achieved more success in matching the empirical variance decomposition by employing time-varying risk aversion and/or time-varying volatility of consumption growth. However, these models must still rely on the assumption very high risk aversion. Overall, we conclude that it remains difficult to justify the observed volatility of stock prices using moderate levels of risk aversion.

The remainder of the paper is organized as follows. Section 2 reviews theoretical variance bounds under the assumption of risk neutrality. Sections 3, 4, and 5 expand the analysis to consider risk aversion in the context of a standard asset pricing model with power utility. Section 6 maps our theoretical results to the empirical framework used by Campbell and Shiller (1988) and others. Section 7 concludes. An appendix provides the details for all

---

6The most notable examples are Campbell and Cochrane (1999) and Bansal and Yaron (2004).
derivations.

## 2 Variance Bounds under Risk Neutrality

Whether or not agents are risk neutral, 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 $p_t$ equals the conditional expectation of $p_t^*$ implies that the variance of $p_t^*$ is an upper bound for 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 $\tilde{p}_t = E(p_t^* | H_t)$, where $\tilde{p}_t$ is the appropriate stock price for an econometrician who has no information useful in forecasting $p_t^*$ other than current and past dividends.\(^7\)

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, current earnings or forecasts of future earnings 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 $I_t$ is a generic characterization of information sets that are at least as fine as $H_t$, but coarser than perfect information about the future, and $J_t$ 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 $Var(\tilde{p}_t)$ is a lower bound for $Var(\overline{p}_t)$, where $\overline{p}_t \equiv E(p_t^* | J_t)$. Thus we have

$$Var(\tilde{p}_t) \leq Var(\overline{p}_t) \leq Var(p_t^*), \tag{1}$$

where the upper and lower variance bounds can be calculated from a univariate model for dividends. The bounds can thus be derived without explicitly specifying the extent of investors’ auxiliary information.

We define the “excess payoff” under information set $I_t$ as

$$\nu_{t+1} \equiv p_{t+1} + d_{t+1} - \beta^{-1} p_t, \tag{2}$$

which represents next-period’s cash value from the stock 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

\(^7\)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 $I_t$. 


to show that the gross risk-free rate equals $\beta^{-1}$. 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

$$\beta \nu_{t+1} + \beta^2 \nu_{t+2} + \beta^3 \nu_{t+3} + \ldots = -p_t + \beta d_{t+1} + \beta^2 d_{t+2} + \beta^3 d_{t+3} + \ldots.$$  \hspace{1cm} (3)

Solving equation (3) for $p^*_t$ and then taking the variance of both sides yields

$$\text{Var}(p^*_t) = \text{Var}(p_t) + \frac{\beta^2}{1 - \beta^2} \text{Var}(\nu_t),$$  \hspace{1cm} (4)

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 $\text{Var}(\nu^*_t) = 0$. Since $\text{Var}(p^*_t) - \text{Var}(p_t) \geq 0$ from equation (1), the above expression establishes that $\text{Var}(\nu_t) \geq 0 = \text{Var}(\nu^*_t)$.

Similarly, we define the excess payoff under information set $H_t$ as $\nu_{t+1} = \hat{p}_{t+1} + d_{t+1} - \beta^{-1} \hat{p}_t$. Following the same methodology as above, we obtain

$$\text{Var}(p^*_t) = \text{Var}(\hat{p}_t) + \frac{\beta^2}{1 - \beta^2} \text{Var}(\nu_t).$$  \hspace{1cm} (5)

Substituting for $\text{Var}(p^*_t)$ from equation (4) into equation (5) and noting that $\text{Var}(\hat{p}_t) - \text{Var}(\nu_t) \geq 0$ from equation (1) establishes that $\text{Var}(\nu_t) \geq \text{Var}(\hat{p}_t)$. 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):

$$\text{Var}(\nu^*_t) = 0 \leq \text{Var}(\hat{p}_t) \leq \text{Var}(\nu_t).$$  \hspace{1cm} (6)

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 payoff variance associated with $H_t$ represents an upper bound for the payoff variance associated with $I_t$.

3 Asset Pricing Model

Equity shares are priced using a frictionless pure exchange model along the lines 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

\[ E \left\{ \sum_{t=0}^{\infty} \beta^t \frac{c_{t}^{1-\alpha} - 1}{1 - \alpha} I_t \right\}, \tag{7} \]

subject to the budget constraint

\[ c_t + p_t s_t = (p_t + d_t) s_{t-1}, \quad c_t, \ s_t > 0, \tag{8} \]

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

\[ p_t = E \left\{ \beta \left( \frac{c_{t+1}}{c_t} \right)^{-\alpha} (p_{t+1} + d_{t+1}) I_t \right\}. \tag{9} \]

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

\[ p_t = E \left\{ \sum_{j=1}^{\infty} M_{t,t+j} d_{t+j} I_t \right\}, \tag{10} \]

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

\[ p_t^* = \sum_{j=1}^{\infty} M_{t,t+j} d_{t+j}. \tag{11} \]

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

4 Volatility of the Price-Dividend Ratio

Dividends, and therefore also the stock price variables \( p_t \) and \( p_t^* \), trend upward, implying that dividend and price variances 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.\(^8\)

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.\(^9\) Current practice, particularly when using a homothetic utility specification like the function (7), 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). This is the procedure we follow here.

To ensure that equilibrium values of intensive variables are stationary, we assume that the growth rate of dividends \(x_t \equiv \log \left(\frac{d_t}{d_{t-1}}\right)\) is governed by the following AR(1) process:

\[
x_{t+1} = \rho x_t + (1 - \rho) \mu + \varepsilon_{t+1}, \quad \varepsilon_{t+j} \sim N \left(0, \sigma^2_\varepsilon\right), \text{ iid, } \quad |\rho| < 1. \tag{12}
\]

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). The geometric random walk model provides a reasonable representation of dividends in the data. However, we do not want to restrict our results to the case where dividend growth rates are iid, so we allow for serially correlated dividend growth \((\rho \neq 0)\).

The specification (12) implies that the unconditional moments of dividend growth are given by

\[
E(x_t) = \mu, \quad \tag{13}
\]

\[
Var(x_t) = \frac{\sigma^2_\varepsilon}{1 - \rho^2}, \quad \tag{14}
\]

\[
Cov(x_{t+j}, x_t) = \rho^j Var(x_t). \quad \tag{15}
\]

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 \(\bar{y}_t \equiv \bar{p}_t/d_t\), 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

---

\(^8\) Otrok, Ravikumar, and Whiteman (2002) document the shifting autocorrelation properties of U.S. consumption growth.

\(^9\) West (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

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

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

\[
y^*_t = \beta \exp [(1 - \alpha) x_{t+1}] (y^*_{t+1} + 1). \tag{18}
\]

The fact that \( \bar{y}_t \) 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

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

Hence, the basic form of the variance bound derived in the earlier literature under risk neutrality, i.e., \( Var(\hat{p}_t) \leq Var(\bar{p}_t) \leq Var(p^*_t) \), 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 obtain 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:

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

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:

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

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

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

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

and \( a_0 \equiv \exp \{ E [\log (\tilde{z}_t)] \} \) is given by

\[
a_0 = \frac{\beta \exp [(1 - \alpha)\mu]}{1 - \beta \exp [(1 - \alpha)\mu + \frac{1}{2} (a_1)^2 \sigma^2_\epsilon]},
\]

provided that \( \beta \exp [(1 - \alpha)\mu + \frac{1}{2} (a_1)^2 \sigma^2_\epsilon] < 1 \).

**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.\(^{10}\) Given the approximate solution for the composite variable \( \tilde{z}_t \), we can recover \( \hat{y}_t \) as follows:

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

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:

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

which in turn can be used to derive an expression for \( \text{Var} (\hat{y}_t) \).\(^{11}\)

From equation (23), the direction of the effect of dividend growth fluctuations on \( \hat{y}_t \) depends on the sign of \( a_1 \), which in turn depends on the values of \( \rho \) and the coefficient of relative risk aversion \( \alpha \). 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 \( \hat{y}_t \), if current dividend growth is high. On the other hand when \( \alpha > 1 \) such

\(^{10}\)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.

\(^{11}\)Given the unconditional mean \( E [\log (\hat{y}_t)] = \log (a_0) + (a_1)^2 \sigma^2_\epsilon / 2 \) and the expression for \( \text{Var} [\log (\hat{y}_t)] \) from equation (24), the unconditional variance of \( \hat{y}_t \) itself can be computed by making use of the following expressions for the mean and variance of the log-normal distribution: \( E (\hat{y}_t) = \exp \{ E [\log (\hat{y}_t)] + \frac{1}{2} \text{Var} [\log (\hat{y}_t)] \} \) and \( \text{Var} (\hat{y}_t) = E (\hat{y}_t)^2 \{ \exp (\text{Var} [\log (\hat{y}_t)]) - 1 \} \).
that $a_1 < 0$, we have $a_1 \rho > 0$. In this case, the expected lower 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 $\widehat{y}_t$ today. All of these effects are reversed when $\rho > 0$.

In the special case of logarithmic utility where $\alpha = 1$, fluctuations in dividend growth do not affect $\log(\widehat{y}_t)$, which is therefore constant. This result obtains 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 $\alpha$ affect the variance of $\log(\widehat{y}_t)$. When $\alpha < 1$, increases in $\alpha$ shrink the magnitude of $a_1$ which moves the variance of $\log(\widehat{y}_t)$ 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 the magnitude of fluctuations in marginal utility exceed the magnitude of fluctuations in inverse consumption, thereby increasing the variance of $\log(\widehat{y}_t)$.

### 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 predict future dividends. We now relax that assumption by allowing investors to see dividends one period ahead, 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:

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

As specified above, $\overline{p}_t$ and $\overline{y}_t$ are the price and price-dividend ratio under $J_t$, while $\widehat{p}_t$ and $\widehat{y}_t$ 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 $\widehat{y}_t = E(\overline{y}_t|H_t)$, which in turn implies $\text{Var}(\widehat{y}_t) \leq \text{Var}(\overline{y}_t)$.

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

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

4.3 Variance under Perfect Foresight

Under perfect foresight, investors are assumed to be able to see all future dividends, not just dividends one period in advance. 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):

$$\log(y^*_t) - E[\log(y^*_t)] \approx (1-\alpha)(x_{t+1} - \mu) + \beta \exp\left[(1-\alpha)\mu\right]\left\{\log(y^*_{t+1}) - E[\log(y^*_t)]\right\}.$$ \hspace{1cm} (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):

$$Var[\log(y^*_t)] = \frac{(1-\alpha)^2}{1-\beta^2 \exp[2(1-\alpha)\mu]} \left\{\frac{1+\rho \beta \exp[(1-\alpha)\mu]}{1-\rho \beta \exp[(1-\alpha)\mu]}\right\} Var(x_t),$$ \hspace{1cm} (29)

which is more complicated than either $Var[\log(\bar{y}_t)]$ from equation (24) or $Var[\log(\bar{y}_t)]$ 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.\textsuperscript{12} We choose parameters to match the mean, standard deviation, and autocorrelation of consumption growth in the data. Using the moment formulas given by equations (13) through (15), our calibration 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 $E[\log(\bar{y}_t)] = 3.18$ 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.\textsuperscript{13} When $\alpha$ exceeds a value of about 3, achieving the target value of $E[\log(\bar{y}_t)]$ 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 $E[\beta(c_{t+1}/c_t)^{-\alpha}]$ remains below unity.\textsuperscript{14}


\footnote{13}Cochrane (1992) employs a similar calibration procedure. For a given discount factor $\beta$, he chooses the risk coefficient $\alpha$ to match the mean price-dividend ratio in the data.

\footnote{14}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

This section presents a quantitative analysis of the volatility of the log price-dividend ratio in the calibrated version of the model described above. 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 \( \log(\hat{y}_t) \) (solid blue line), \( \log(y_t) \) (dotted grey line), and \( \log(y_t^*) \) (dashed red line).

The standard deviation of \( \log(\hat{y}_t) \) 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. The model-predicted volatility for \( \log(\hat{y}_t) \) 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. The model-predicted volatility for \( \log(\hat{y}_t) \), which is based on the assumption that investors see dividends one period ahead, is noticeably higher than the volatility of \( \log(\hat{y}_t) \), 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.

A finding of excess volatility requires \( \text{Var}[\log(y_t^{\text{us}})] > \text{Var}[\log(y_t^*)] \). Figure 1 shows that excess volatility prevails for \( \alpha < 5 \). In contrast, for \( \alpha > 5 \), we have \( \text{Var}[\log(y_t^{\text{us}})] < \text{Var}[\log(y_t^*)] \), so we cannot make a finding of excess volatility. The interpretation is that the volatility of \( \log(y_t^{\text{us}}) \) is consistent with the present-value model if real-world investors are risk averse and have access to very good information about future dividend growth. 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 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 price volatility can be connected either with unconditional variance measures, corresponding to a long-run interpretation of volatility, or with conditional variance measures, corresponding to a short-run concept. We also noted that, based on earlier research assuming risk neutrality, the present-value model has implications

\(^{15}\text{The standard deviation of the U.S. price dividend ratio in levels (as opposed to logarithms) is 13.8, with a corresponding mean value of 26.6.}\)
for both measures of volatility. Specifically, the variance-bounds tests involve determining whether the joint restrictions implied by the present-value model for both types of volatility measures are satisfied. So far we have concentrated on bounds for unconditional price volatility as embodied in $\text{Var} \left[ \log (y_{t}) \right]$ and $\text{Var} \left[ \log (y_{t}^{*}) \right]$.

We now turn to measures of short-run price volatility. There are several ways to gauge short-run volatility: the variance of one-period excess payoffs, the unconditional variance of the log price change, or the unconditional variance of the rate of return. Since these measures are highly correlated, it does not matter much for the substantive results which measure is used.\(^{16}\) It turns out that, just as the variance of one-period excess payoffs is the most convenient measure of short-run volatility when dividends are governed by a stationary linear process, the variance of log returns is the most convenient in the setting considered here.

The gross rates of return on equity under the various information assumptions can be written as

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

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

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

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

\[
R_{t+1}^{*} = \frac{p_{t+1}^{*} + d_{t+1}}{p_{t}^{*}} = \exp(x_{t+1}) \left( \frac{y_{t+1}^{*} + 1}{y_{t}^{*}} \right)
\]

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

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

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

which follows directly from equation (20). To obtain a similar return expression for information set $J_{t}$, we define the composite variable $\bar{z}_{t+1} \equiv \beta \exp \left[ (1 - \alpha) x_{t+1} \right] (\bar{y}_{t+1} + 1)$ and use\(^{16}\) Over the period 1871 to 2008, the correlation coefficient between log real equity returns and log real price changes in U.S. data is 0.994. LeRoy (1984, p. 186) shows that the conditional price variance is numerically very close to the unconditional variance of price changes in a calibrated asset pricing model.
this definitional relationship and the corresponding equilibrium condition \( \bar{y}_t = E(\bar{z}_{t+1}|J_t) \) to eliminate \( \bar{y}_t \) and \( \bar{y}_{t+1} + 1 \) from equation (31). In the expression for \( R^*_t \), we have substituted in \( (y^*_t + 1)/y^*_t = \beta^{-1} \exp[-(1-\alpha)x_{t+1}] \) from the nonlinear law of motion (18). Notice that the three return measures differ only by the terms \( \bar{z}_{t+1}/E(\bar{z}_{t+1}|H_t) \) 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 under risk neutrality, which also differed only in terms of the size of the investor’s forecast errors.

In Appendix A.2, we show that the approximate law of motion for \( \log(\hat{R}_{t+1}) \) is

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

(34)

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 \( \log(\hat{R}_{t+1}) \) is

\[
\log(\hat{R}_{t+1}) - E[\log(\hat{R}_{t+1})] = n_1 (x_{t+2} - \mu) + (1 - a_1) (x_{t+1} - \mu),
\]

(35)

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 \( \log(R^*_t) \) is

\[
\log \left( R^*_t \right) - E \left[ \log \left( R^*_t \right) \right] = \alpha (x_{t+1} - \mu).
\]

(36)

Given the above laws of motion for log returns, it is straightforward to compute the following unconditional variances:

\[
Var[\log(\hat{R}_{t+1})] = \alpha^2 Var (x_t) + a_1 [a_1 + 2\alpha] \sigma^2_{\varepsilon},
\]

(37)

\[
Var \left[ \log(\hat{R}_{t+1}) \right] = [(n_1)^2 + (1 - a_1)^2 + 2n_1 (1 - a_1) \rho] Var (x_t),
\]

(38)

\[
Var \left[ \log \left( R^*_t \right) \right] = \alpha^2 Var (x_t).
\]

(39)

### 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 \( Var (x_t) = \sigma^2_{\varepsilon} \). The variance expressions imply the following inequality:

\[
\underbrace{Var \left[ \log \left( R^*_t \right) \right]}_{=0} \leq \underbrace{Var \left[ \log(\hat{R}_{t+1}) \right]}_{= (n_1)^2 \sigma^2_{\varepsilon}} \leq \underbrace{Var[\log(\hat{R}_{t+1})]}_{=\sigma^2_{\varepsilon}}, \quad \text{when } \alpha = \rho = 0,
\]

(40)

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. This finding agrees with that of LeRoy and Parke (1992) in a similar (but not identical) setting.

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

\[
\begin{align*}
\text{Var} \left[ \log(R_{t+1}^*) \right] \leq \text{Var} \left[ \log(\hat{R}_{t+1}) \right] \leq \text{Var} \left[ \log(\hat{R}_{t+1}) \right],
\end{align*}
\]

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

\[
\begin{align*}
\text{Var} \left[ \log(R_{t+1}^*) \right] &= \text{Var} \left[ \log(\hat{R}_{t+1}) \right] = \text{Var} \left[ \log(\hat{R}_{t+1}) \right], \quad \text{when} \ \alpha = 1,
\end{align*}
\]

for every specification of \( I_t \). Since the price-dividend ratio is constant under log utility regardless of the representative investor’s information, 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 \( \text{Var}[\log(R_{t+1}^*)] \) can also occur when \( a_1 + 2\alpha = 0 \). We will verify below that at least for some parameter specifications there exists a positive value of \( \alpha \) that satisfies this equation. The critical value of \( \alpha \) defines a crossing point at which the size ordering between \( \text{Var}[\log(\hat{R}_{t+1})] \) and \( \text{Var}[\log(R_{t+1}^*)] \) reverses. Further, it turns out that at the critical value of \( \alpha \) we have \( \text{Var}[\log(\hat{R}_{t+1})] > \text{Var}[\log(\hat{R}_{t+1})] \) and \( \text{Var}[\log(R_{t+1}^*)] \) is the return variance based on \( J_t = H_t \cup d_{t+1} \). We therefore conclude that \( \text{Var}[\log(\hat{R}_{t+1})] \) and \( \text{Var}[\log(R_{t+1}^*)] \) cannot be bounds for return volatility. This result should not be surprising. Unlike the situation with prices, 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.

5.2 Quantitative Analysis

Solving for the critical value of \( \alpha \) where \( \text{Var}[\log(\hat{R}_{t+1})] = \text{Var}[\log(R_{t+1}^*)] \) 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 \( \exp[ (1 - \alpha) \mu + (a_1)^2 \sigma^2 / 2 ] \simeq 1 \) 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 value for \( \alpha \) for
which $\text{Var}[\log(\hat{R}_{t+1})] = \text{Var} [\log(R^*_t)]$. This second value is $\alpha \simeq 1/(2\rho \beta - 1)$. Positivity of $\alpha$ requires that the model parameters satisfy $\rho \beta > 0.5$. This counterexample establishes that $\text{Var} [\log(R^*_t)]$ 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 $\alpha$.

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 $\log(R^*_t)$ 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 predict dividends either one period or an infinite number of periods in the future 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 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. Incidentally, it is surprising that the dependence of return volatility on risk aversion is about the same whether investors can see ahead one period or an infinite number of periods.

In the bottom panel, we set $\rho = 0.7$ and recalibrate the value of $\sigma_\epsilon$ to maintain the same standard deviation of consumption growth as in the top panel. This calibration is extremely unrealistic empirically; we consider it only to illustrate the point made above that for general parameter values the extreme specifications of investors’ information do not define bounds on return volatility. In this case, the model parameters satisfy $\rho \beta > 0.5$ so the two return volatilities are equal not only when $\alpha = 1$, but also when $\alpha \simeq 1/(2\rho \beta - 1) = 2.9$. As $\alpha$ crosses the values 1 and 2.9, the direction of the variance inequality comparing the volatility of $\log(\hat{R}_{t+1})$ to that of $\log(R^*_t)$ reverses direction. As observed above, such reversals demonstrate that $\text{Var} [\log(R^*_t)]$ cannot be a lower bound for all $\alpha$.\footnote{Lansing (2011) shows that a similar variance inequality involving price changes (rather than returns) can also be reversed, depending on parameter values. Moreover, he shows that the price-change variance bounds derived by Engel (2005) for the case of risk-neutrality and “cum-dividend” stock prices do not extend to the case of ex-dividend stock prices.}

6 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 finding of excess volatility in the original variance-bounds literature, where risk neutrality was 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 connection between risk aversion and stock price volatility. Its use up to now in the finance literature has been to determine the relative empirical importance of dividend growth, risk free interest rates, and excess returns in explaining the volatility of real-world stock prices relative to dividends. However, since the return identity is 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 sources noted above for comparison with the results obtained from real-world data.

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

$$\log(\hat{R}_{t+1}) \equiv \log (\hat{y}_{t+1} + 1) + x_{t+1} - \log (\hat{y}_t), \quad (43)$$

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 (43) 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:

$$\log (\hat{y}_t) \simeq \frac{\kappa_0}{1 - \kappa_1} + \sum_{j=0}^{\infty} (\kappa_1)^j \left[ x_{t+1+j} - \log(\hat{R}_{t+1+j}) \right], \quad (44)$$

assuming that the summation converges. The convergence assumption, which implies that the determinants of the log price-dividend ratio are not pushed off to the infinite future, is satisfied in our model. It follows 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 $\log(\bar{y}_t)$ and $\log(y^*_t)$, under information sets $J_t$ and perfect foresight, respectively.

The variables in the approximate identity (44) 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

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

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

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

(45)

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 log excess return on equity, given by \( \log(\hat{R}_{t+1+j}^f/\hat{R}_{t+1+j}^f) \). Analogous decompositions can be derived for \( \text{Var} \left[ \log (\hat{y}_t) \right] \) and \( \text{Var} \left[ \log (y_t^e) \right] \) which involve covariance terms with \( \log(\hat{R}_{t+1+j}) \) and \( \log(R_{t+1+j}^e) \), 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 a regressor.

For our model, the approximate laws of motion for the log equity return under each information set are given by equations (34) through (36). In the appendix, we show that the corresponding laws of motion for the log risk-free rate are given by

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

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

(46)

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

(47)

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

(48)

Using the approximate laws of motion for the relevant variables, we can analytically compute the three covariance terms in the applicable version of equation (45) for each information set. Details are provided in the appendix. The results of the theoretical variance decomposi-
tion are as follows:

\[ \text{Var} \left[ \log (\tilde{y}_t) \right] = \frac{a_1 \rho^2 \text{Var}(x_t)}{1 - \rho \kappa_1} - \frac{a_1 \rho^2 \text{Var}(x_t)}{1 - \rho \kappa_1} - 0, \quad (49) \]

\[ \text{Var} \left[ \log (y_t) \right] = \frac{a_1 \text{Var}(x_t)}{1 - \rho \kappa_1} - \frac{a_1 \text{Var}(x_t)}{1 - \rho \kappa_1} - \left[ a_1 (1 - \alpha) - (a_1)^2 \right] \text{Var}(x_t), \quad (50) \]

\[ \text{Var} \left[ \log (y_t) \right] = \frac{(1 - \alpha) (1 + \rho \kappa_1) \text{Var}(x_t)}{1 - (\kappa_1)^2 (1 - \rho \kappa_1)} - \frac{(1 - \alpha) (1 + \rho \kappa_1) \text{Var}(x_t)}{1 - (\kappa_1)^2 (1 - \rho \kappa_1)} - 0, \quad (51) \]

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 Taylor-series coefficient \( \kappa_1 \) in the approximate return identity 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 (49) and (51) 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 derived in the appendix and reproduced below:

\[ \log(\tilde{R}_{t+1}) - \log(\tilde{R}^f_{t+1}) = (a_1 + \alpha) \varepsilon_{t+1} - \frac{1}{2} [(a_1)^2 - \sigma^2] \varepsilon^2, \quad (52) \]

\[ \log(R_{t+1}) - \log(R^f_{t+1}) = (1 - \alpha - a_1 + \rho n_1) (x_{t+1} - \mu) + n_1 \varepsilon_{t+2}, \quad (53) \]

\[ \log(R^*_t) - \log(R^*_t) = 0. \quad (54) \]

Equation (52) shows that excess returns are \textit{iid} under information set \( H_t \), while equation (54) shows that excess returns are identically zero under perfect foresight.\( ^{18} \) In both cases, the covariance between future excess returns and the log price-dividend ratio at time \( t \) is zero. In contrast, equation (53) shows that excess returns under information set \( J_t \) are not \textit{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 \textit{iid} such that the variance contribution from excess returns will be zero. In this case, the third term in equation (50) will also be zero. But even when \( \rho \neq 0 \), the contribution from the third term turns out to be numerically very small for information set \( J_t \).

The theoretical variance decomposition in equations (49) through (51) 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>$\frac{1}{a_1(1-\rho)}$</td>
<td>$-\frac{\alpha}{a_1(1-\rho)}$</td>
<td>$-\left[\frac{(1-\alpha)}{a_1(1-\rho)} - 1\right]$</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. 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, as discussed further below.

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 increasing in the risk aversion coefficient $\alpha$. The intuition for this result is straightforward. The risk aversion coefficient influences the volatility of the stochastic discount factor, which in turn influences the volatility of the risk-free rate. Under risk neutrality with $\alpha = 0$, equations
(46) through (48) show that the risk-free rate is constant and thus contributes nothing to the variance of the log price-dividend ratio. But as $\alpha$ increases, a larger fraction of the variance is attributable to this source. 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 changing forecasts of future dividend/consumption growth, $125\%$ is attributable to changing forecasts of future risk free rates, and $0\%$ is attributable to changing forecasts of future excess returns.$^{19}$ 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 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 changing forecasts of future dividend growth, while $138\%$ is attributable to changing forecasts of future equity 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. 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 the 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 those in the data. In the data, more than $100\%$ of the variation in the log price-dividend ratio is attributable to future excess returns, i.e., the third term in equation (45), while almost nothing can be attributed to 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 future risk-free rates, while nothing

---

$^{19}$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%.
(or almost nothing, depending on the information set) is attributable to 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.

7 Conclusion

Our first conclusion is that for low levels of risk aversion, the volatility of observed price-dividend ratios in U.S. data greatly exceeds the upper bound implied by the present-value model. Thus we reproduce the result found in the earlier variance-bounds literature. However, the finding of excess volatility disappears under the assumption that investors have moderately high risk aversion; 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 perfect foresight. In other words, the volatility of stock prices is about as one would expect under the assumption that investors can forecast future dividends accurately into the distant future. To be sure, the assumption that investors have such foresight may be viewed as implausible. If one includes in the null hypothesis the more realistic assumption that investors can forecast dividend growth at most only a few years into the future, then actual stock price volatility appears excessive for any level of risk aversion.

The foregoing conclusion is reminiscent of that of Mehra and Prescott (1985), who found that they could not explain the equity premium in a power-utility model with low or moderate levels of risk aversion, but could do so with extremely high 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 long-run U.S. data.

We also demonstrate that the bounds on excess payoff variance derived under the assumptions of risk neutrality and a stationary linear process for dividends do not extend to analogous bounds on log equity returns in an environment with risk aversion and exponentially growing dividends, except in some special cases.

20Mehra and Prescott (1985) argue that risk aversion coefficients below 10 are plausible.
Further analysis of the power utility model raises doubts about whether observed stock price 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 (model-free) analysis 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 forecastable 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. From the tremendous amount of research in the area it appears that such a model must incorporate either time-varying risk aversion and/or time-varying volatility of consumption growth. In this way, excess returns can be made to exhibit significant persistence and volatility, in contrast to equations (52) through (54). For example, Campbell and Cochrane (1999) introduce 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 set equal to 10.21 It therefore remains a challenge for theoretical asset pricing models to explain the volatility of observed stock prices relative to dividends, as well as other features of the data, using reasonable values for risk aversion.

---

21See 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 $\tilde{z}_t$ specified in Proposition 1 and taking the conditional expectation implied by the information set $H_t$ yields

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

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

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

$$\simeq \log (a_0) + a_1 (x_t - \mu),$$  \hfill (A.2)

where the Taylor-series coefficients $a_0$ and $a_1$ are given by

$$\log (a_0) = F(\mu) = \log (\beta) + (1 - \alpha) \mu + \log \left\{ a_0 \exp \left[ \frac{1}{2} (a_1)^2 \sigma^2 \right] + 1 \right\}$$  \hfill (A.3)

$$a_1 = \frac{\partial F}{\partial x_t}\bigg|_{\mu} = 1 - \alpha + \frac{\rho a_1 a_0 \exp \left[ \frac{1}{2} (a_1)^2 \sigma^2 \right]}{a_0 \exp \left[ \frac{1}{2} (a_1)^2 \sigma^2 \right] + 1}. \hfill (A.4)$$

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

$$a_0 = \exp \left\{ E[\log (\tilde{z}_t)] \right\} = \frac{\beta \exp [(1 - \alpha) \mu]}{1 - \beta \exp \left[ (1 - \alpha) \mu + \frac{1}{2} (a_1)^2 \sigma^2 \right]},$$  \hfill (A.5)

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

$$a_1 = 1 - \alpha + \rho a_1 \beta \exp \left[ (1 - \alpha) \mu + \frac{1}{2} (a_1)^2 \sigma^2 \right].$$ \hfill (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 \( \log (\hat{y}_t) \) in equation (23) yields

\[
E [\log (\hat{y}_t)] = \log (a_0) + \frac{1}{2} (a_1)^2 \sigma^2.
\]  

(A.7)

We then have

\[
\log (\hat{y}_t) - E [\log (\hat{y}_t)] = a_1 \rho (x_t - \mu),
\]  

(A.8)

which in turn implies

\[
Var [\log (\hat{y}_t)] = (a_1 \rho)^2 Var (x_t).
\]  

(A.9)

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

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

(A.10)

Substituting in \( E(\hat{z}_{t+1} | H_t) \) 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 \( \log (\tilde{R}_{t+1}) \) yields

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

(A.11)

We then have

\[
\log (\tilde{R}_{t+1}) - E[\log (\tilde{R}_{t+1})] = \alpha (x_{t+1} - \mu) + a_1 \varepsilon_{t+1},
\]  

(A.12)

which in turns implies

\[
Var[\log (\tilde{R}_{t+1})] = \alpha^2 Var (x_t) + (a_1)^2 \sigma^2 + 2 \alpha a_1 Cov (x_{t+1}, \varepsilon_{t+1}) = \sigma^2.
\]  

(A.13)

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

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

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

\[
= - \log (\beta) + \alpha [px_t + (1 - \mu) \mu] - \frac{1}{2} \alpha^2 \sigma^2,
\]  

(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 \( \log (\tilde{R}_{t+1}^f) \) and then subtracting the unconditional mean from equation (A.14) yields the law of motion (46).
A.3 Variance Decomposition

For information set $H_t$, we have

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

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

Using the law of motion for $\log (b_y t)$ 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 (45):

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

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

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

Similarly, using the law of motion for $\log(\tilde{R}_{t+1}^f)$ given by (46) we can compute the following covariance which is the second term in equation (45):

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

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

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

The law of motion for excess returns (52) shows that excess returns are iid under information set $H_t$. Hence, the third covariance term in equation (45) is identically zero. Dividing both sides of equation (45) by $Var [\log (\tilde{y}_t)]$ 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

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

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 $\bar{y}_t \equiv \bar{p}_t/d_t$:

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

$$\bar{y}_t = \hat{z}_{t+1}, \quad (B.2)$$

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 $\bar{y}_t$:

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

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

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

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

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

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

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

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

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

$$= -\log (\beta) + \alpha \mu \quad (B.7)$$

We then have

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

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

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

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

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

$$= -\log (\beta) + \alpha x_{t+1}, \quad (B.9)$$

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 $\log(\bar{R}^f_{t+1})$ and then subtracting the unconditional mean from equation (B.9) yields the law of motion (47).

**B.3 Variance Decomposition**

For information set $J_t$, we have

$$\kappa_1 = \frac{\exp \left[ E \log (\bar{y}_t) \right]}{1 + \exp \left[ E \log (\bar{y}_t) \right]} = \frac{a_0}{1 + \bar{a}_0} = \frac{\beta \exp ((1-\alpha)\mu)}{1 + \beta \exp ((1-\alpha)\mu) - \beta \exp ((1-\alpha)\mu + \frac{1}{2}(a_1)\sigma^2)}, \quad (B.10)$$

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:

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

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

\[
= \frac{a_1 \text{Var} (x_t)}{1 - \rho \kappa_1}.
\]  

(B.11)

Similarly, using the law of motion for \( \log(\bar{R}_{t+1}^f) \) given by (47) we can compute the following covariance which is the second term in the variance decomposition:

\[
- \text{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\},
\]

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

\[
= \frac{\alpha a_1 \text{Var} (x_t)}{1 - \rho \kappa_1}.
\]

(B.12)

Using the law of motion for excess returns (53), the third term in the variance decomposition is given by

\[
- \text{Cov} \left[ \log (\bar{y}_t), \sum_{j=0}^{\infty} (\kappa_1)^j \log \left( \frac{\bar{R}_{t+1+j}^f}{\bar{R}_{t+1}^f} \right) \right] = -E \left\{ a_1 (x_{t+1} - \mu) (1 - \alpha - a_1 + \rho n_1) (x_{t+1} - \mu)
\right.

+ \kappa_1 a_1 (x_{t+1} - \mu) (1 - \alpha - a_1 + \rho n_1) (x_{t+2} - \mu) + \ldots \right\},
\]

\[
= a_1 (1 - \alpha - a_1 + \rho n_1) \text{Var} (x_t) \left\{ 1 + \rho \kappa_1 + (\rho \kappa_1)^2 + \ldots \right\},
\]

\[
= \frac{a_1 (1 - \alpha - a_1 + \rho n_1) \text{Var} (x_t)}{1 - \rho \kappa_1},
\]

\[
= \frac{a_1 [1 - \alpha - a_1 (1 - \rho \kappa_1)] \text{Var} (x_t)}{1 - \rho \kappa_1}.
\]

(B.13)

where we note that \( n_1 = \kappa_1 a_1 \).
Dividing both sides of the variance decomposition by \( \text{Var} \left[ \log (y_t) \right] \) from equation (B.4) and then substituting in the appropriate moments yields
\[
1 = \frac{1}{a_1 (1 - \rho \kappa_1)} - \frac{\alpha}{a_1 (1 - \rho \kappa_1)} - \left[ \frac{(1 - \alpha)}{a_1 (1 - \rho \kappa_1)} - 1 \right],
\] (B.14)
where equations (A.6) and (B.10) imply \((1 - \rho \kappa_1) \simeq (1 - \alpha) / a_1\) only when \(\rho \simeq 0\).

C Solution under Perfect Foresight

C.1 Log-linearized Law of Motion

Taking logarithms of the nonlinear law of motion (18) yields
\[
\log (y^*_t) = G \left[ x_{t+1}, \log (y^*_t) \right] = \log (\beta) + (1 - \alpha) x_{t+1} + \log \{ \exp \left[ \log (y^*_t) \right] + 1 \}
\]
\[
\simeq \log (b_0) + b_1 (x_{t+1} - \mu) + b_2 \left[ \log (y^*_{t+1}) - \log (b_0) \right],
\]
(C.1)
where the Taylor-series coefficients \(b_0, b_1,\) and \(b_2\) are given by
\[
\log (b_0) = G [\mu, \log (b_0)] = \log (\beta) + (1 - \alpha) \mu + \log [b_0 + 1],
\]
(C.2)
\[
b_1 = \frac{\partial G}{\partial x_t} \bigg|_{\mu, \log (b_0)} = 1 - \alpha,
\]
(C.3)
\[
b_2 = \frac{\partial G}{\partial \log (y^*_{t+1})} \bigg|_{\mu, \log (b_0)} = \frac{b_0}{b_0 + 1}.
\]
(C.4)

Solving equation (C.2) for the unconditional mean \(b_0\) yields
\[
b_0 = \exp \left\{ E \left[ \log (y^*_t) \right] \right\} = \frac{\beta \exp \left[ (1 - \alpha) \mu \right]}{1 - \beta \exp \left[ (1 - \alpha) \mu \right]},
\]
(C.5)
which can be substituted into equation (C.4) to obtain the following expression for \(b_2\) :
\[
b_2 = \beta \exp \left[ (1 - \alpha) \mu \right].
\]
(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

\[ \text{Var} \left[ \log (y^*_t) \right] = \frac{(1 - \alpha)^2 \text{Var} (x_t) + 2 (1 - \alpha) \beta \exp [(1 - \alpha) \mu] \text{Cov} \left[ \log (y^*_t), x_t \right]}{1 - \beta^2 \exp [2 (1 - \alpha) \mu]} . \]  

(C.7)

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

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

(C.8)

\[ \text{Cov} \left[ \log (y^*_{t+1}), x_t \right] = (1 - \alpha) \text{Cov} \left[ x_{t+1}, x_{t-1} \right] = \rho \text{Cov}(x_t, x_{t-1}) + \beta \exp [(1 - \alpha) \mu] \text{Cov} \left[ \log (y^*_{t+2}), x_t \right] , \]  

(C.9)

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

\[ \text{Cov} \left[ \log (y^*_t), x_t \right] = (1 - \alpha) \text{Cov} \left[ x_{t}, x_{t-1} \right] \sum_{j=0}^{\infty} \{ \rho \beta \exp [(1 - \alpha) \mu] \}^j \]  

\[ = \frac{(1 - \alpha) \text{Cov} \left[ x_{t}, x_{t-1} \right]}{1 - \rho \beta \exp [(1 - \alpha) \mu]} = \frac{(1 - \alpha) \rho \text{Var} (x)}{1 - \rho \beta \exp [(1 - \alpha) \mu]} , \]  

(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

\[ E \left[ \log (R^*_{t+1}) \right] = - \log (\beta) + \alpha \mu . \]  

(C.12)

We then have

\[ \log (R^*_{t+1}) - E \left[ \log (R^*_{t+1}) \right] = \alpha (x_{t+1} - \mu) , \]  

(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

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

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

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

(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 (48).

C.3 Variance Decomposition

Under perfect foresight, we have

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

(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:

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

\[
+ \kappa_1 (1 - \alpha) (x_{t+1} - \mu) (x_{t+2} - \mu) + \kappa_2 b_2 \log (y_{t+1}^*/b_0) (x_{t+2} - \mu)
\]

\[
+ (\kappa_1)^2 (1 - \alpha) (x_{t+1} - \mu) (x_{t+3} - \mu)
\]

\[
+ (\kappa_1)^2 b_2 \log (y_{t+1}^*/b_0) (x_{t+3} - \mu) + \ldots \right\}
\]

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

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

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

\[
+ \kappa_1 \Cov [\log (y_t^*), x_t] \left\{ 1 + (\kappa_1)^2 + (\kappa_1)^4 + \ldots \right\}
\]

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

\[
= \frac{(1 - \alpha) (1 + \rho \kappa_1) \Var (x_t)}{[1 - (\kappa_1)^2] (1 - \rho \kappa_1)},
\]

(C.16)
where we have substituted in $Cov[\log(y_t^*), x_t]$ from equation (C.10).

Following the same methodology and using the law of motion for $\log(R_{t+1}^f)$ given by (48), we can compute the following covariance which is the second term in the variance decomposition:

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

The law of motion for excess returns (54) 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 $Var[\log(y_t^*)]$ 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


35


Figure 1: The log price-dividend ratio in U.S. data exhibits excess volatility for $\alpha < 5$. 
Figure 2: The present-value model does not impose bounds on returns in general settings involving risk aversion.