Revisiting Asset Pricing Anomalies in an Exchange Economy*

Christine A. Parlour† Richard Stanton‡ Johan Walden§

October 16, 2009

Abstract

Several well-known asset pricing anomalies arise when simple endowment economies are calibrated to real data. We show that many of these anomalies are largely mitigated, and even disappear, if we endow the representative agent with an arbitrarily small minimum consumption level. We illustrate this point in a standard one-tree Lucas exchange economy with power utility and lognormal consumption growth. Insuring the agent's consumption allows us to solve the model for parameter values where the standard model is not defined. For such parameter values, disasters are much more important for the representative investor, and the equity premium therefore higher, even though the insurance makes consumption *less* risky. Our model yields reasonable risk premia, Sharpe ratios and discount rates; excess price volatility; a high market price-dividend ratio; and an upward-sloping term structure. Technically, our model leads to nonlinear price functions and to price dynamics quite different from those in the standard model. The model is tractable, and we derive closed-form solutions for all variables of interest. We also establish that the results can be generalized.

*We thank Frederico Belo, Jonathan Berk, Nicole Branger, Nicolae Gârleanu, Martin Lettau, Mark Rubinstein and seminar participants at Uppsala University, the 2009 EFA and the Berkeley Lunch Seminar for helpful comments. The explicit characterization of the term structure presented in Section 3.2 first appeared in an unpublished working paper, "The Long-Term Discount Rate," by the same authors.

†Haas School of Business, U.C. Berkeley, parlour@haas.berkeley.edu.

‡Haas School of Business, U.C. Berkeley, stanton@haas.berkeley.edu.

§Haas School of Business, U.C. Berkeley, walden@haas.berkeley.edu.

1 Introduction

When a standard one-tree consumption-based exchange economy, with Brownian log-consumption growth and a representative investor with power utility, is calibrated to data, three significant anomalies arise.[1] First is the *Equity premium puzzle*, famously posed by Mehra and Prescott (1985): For reasonable values of the risk-aversion coefficient, the implied equity premium is too low. Second is the *Risk-free rate puzzle* (see Weil (1989)): If risk aversion is chosen to match the equity premium, then the discount rate is implausible. Third is the *Excess-volatility puzzle* (see LeRoy and Porter (1981), and Shiller (1981)): Price volatility in the standard model is the same as dividend volatility and consumption volatility; in reality, however, price volatility is many times higher than both consumption and dividend volatility. As summarized in LeRoy (2006),

> "*The conclusion that appears to follow from the equity premium and price volatility puzzles is that, for whatever reason, prices of financial assets do not behave as the theory of consumption-based asset pricing predicts.*"

We revisit the standard model and simply introduce an arbitrarily small risk-free consumption stream. We then show that this minor modification largely mitigates the puzzles. We call this the *Minimum Consumption* (MC) economy. Our results are based on the observation that for some parameter values, beyond what we dub the *breakpoint*, the risky tree in the standard model is so risky that the representative investor's expected utility is negative infinity, and the risk premium is therefore not even defined. With a lower bound on consumption, expected utility remains finite, though it is still strongly affected by low consumption states in this parameter region. As a result we obtain a much higher risk premium than in the standard model. Indeed, for low growth rates and personal discount factors, the risk premium in our model for these parameter values approaches $\gamma^2\sigma^2$ instead of the $\gamma\sigma^2$ produced by the standard model. Interestingly, the consumption process in our economy, with probability one, looks indistinguishable from the standard one-tree model in the long run. Empirically, it would therefore be impossible to distinguish the consumption process in the MC model from that in the standard model, even though the differences in asset pricing are huge.

For simplicity, we implement the model in a two-trees framework (see Cochrane, Longstaff, and Santa-Clara (2008)), with one risky and one risk-free tree. This makes the analysis tractable, and we obtain closed-form solutions for all variables of interest, though we also show that the effect of minimum consumption levels extends to broader classes of model.

[1]Though commonly referred to as a "Lucas" model, the first order conditions for this economy and associated stochastic discount factor, $\rho\frac{U'(C_{t+1})}{U'(C_t)}$, were first derived by Rubinstein (1976), and later used by Lucas (1978).

In a simple calibration of the MC model, we show that to obtain a market risk premium of 5% requires a risk aversion coefficient of only $\gamma = 12.2$, compared with the $\gamma = 31$ needed by the standard model.[2] We also show that, in stark contrast to the standard model, the long-term discount rate in our model is *independent* of risk aversion. In the calibration, we get a long rate of 2.4%, so there is no risk-free rate puzzle at the long end of the yield curve. The short rate is -2.8%, which is somewhat low, but far above the -58% implied by the standard model with the same parameters; moreover, instead of the flat term structure in the standard model, we typically get an upward-sloping term structure. Price volatility is also higher than in the standard model: Our calibration yields a price volatility of 10.3%, compared with a consumption volatility of 4%. Finally, our calibration produces a reasonable market Sharpe ratio of 0.49.

Central to our analysis is the existence of a risk-free consumption stream. There are many plausible economic frameworks that give rise to such a sector; we posit two. First, in an economy with technology shocks, if there is enough "memory" in the economy, it is natural to assume that production levels can never fall below some threshold. Similarly, a lower bound on consumption can be interpreted as subsistence farming or consumption.[3] Second, bonds may not be in zero net supply. The assumption that bonds are in zero net supply is consistent with an infinitely lived representative agent in an economy absent any frictions. In particular, any bonds that she issues, she also consumes. By contrast, in a world with finitely lived investors, or with frictions, it may be possible for the current generation to borrow against the consumption of future generations, leading to a positive supply of bonds and risk-free consumption for the current generation over a significant time period. Indeed, in any economy in which Ricardian equivalence fails, government bonds can be in positive net supply.[4]

Intuitively, the existence of a minimum consumption level lowers the value to the representative consumer of claims that pay off in states when her risky consumption is low. The representative consumer weighs two factors when evaluating a claim that pays off when her other consumption is low: first, her current level of consumption, and second, the difference between current marginal utility and marginal utility when the claim pays off. The first

[2]In the standard model, $\gamma = 12.2$ leads to a risk premium of only 2%.

[3]If a cataclysmic event such as a nuclear war occurred, a subsistence level of consumption might not exist. However, since it is also unlikely that financial assets would survive, we restrict our attention to states of the world in which no such event occurs. The only modification of the model is that the representative investor has a higher effective personal discount discount rate in the presence of such events (similar to the increased discount rate in the portfolio problem of an investor with finite, stochastic life length, compared with an infinitely lived one).

[4]In the extreme case, if the representative investor does not care at all about consumption after a certain date, he will take the opportunity to transfer risk-free consumption from beyond that date, if feasible. The economy then behaves like one with a finite horizon and a minimum consumption level. (Our results also hold for long but finite horizons; see Section 4.5).

factor is important because it affects how far into the future she will consume the claim. A higher current consumption level decreases the value of this claim by increasing the time until its payoff (because the personal discount rate is positive). However, a higher current consumption level also increases the relative difference between current marginal utility and marginal utility at payoff. In the paper, we show that the relative importance of these two factors changes drastically when passing the breakpoint. In the region in which the standard model is defined, the first effect dominates the second, so for high consumption levels the price of a low-consumption claim is negligible. Beyond the breakpoint, however, the second effect dominates the first, and the claim becomes more and more valuable, the higher the consumption level. In the standard model, the price of such a claim is infinite, which is why the standard model is not defined beyond the breakpoint. By contrast, in the MC economy, the minimum consumption level leads to a finite, albeit high, price for the claim.

From the previous discussion, it follows that there is a strong connection between our analysis and any that focuses on the importance of low-consumption states; we are certainly not the first to consider the effect of extreme events in consumption based asset pricing. Notably, Barro (2005) (following Rietz (1988)) uses the absence of consumption insurance to generate empirically reasonable equity premia. These approaches consider the effect of catastrophic risk, either actual or suspected, on agents' valuation for risky bets. In a similar spirit, Weitzman (1998, 2001, 2007) argues that parameter uncertainty, through an increased subjective probability for low states, significantly changes the long-term discount rate and equity premium. Our innovation is to incorporate small consumption insurance against the low states, thereby *lowering* the risk. This allows us to analyze asset pricing properties for parameters which are typically ignored; we explore this idea in more detail when we consider the robustness of our model.

A vast literature has suggested other solutions to the classic puzzles, usually based on significant modifications of the standard model. We cannot do justice to this literature here, but mention a few examples. To solve the equity premium puzzle, some researchers have explored preference specifications that make the stochastic discount factor (SDF) more volatile. Abel (1990) introduced catching-up-with-the-Joneses preferences, while Campbell and Cochrane (1999) suggested that consumers form habits. Others have investigated rational bubbles as a potential solution to the excess-volatility puzzle (see, for example, Blanchard (1979), Blanchard and Watson (1982), Froot and Obstfeld (1991)). With rational bubbles, prices are highly nonlinear functions of dividends, leading to a higher price volatility. In our model, the market price of equity is a convex function of consumption, which mechanically leads to a higher risk premium and price volatility. This is similar to the price behavior in, for example, Abel (1990), and Froot and Obstfeld (1991). In contrast to these models, however, we make minimal modifications to the standard model; preferences are the same and there are no bubbles in the MC economy. The only difference is the addition of an arbitrarily small additional consumption stream.

The rest of the paper is structured as follows. We proceed by laying out the MC model in Section 2, and study when the differences between this and the standard economy are important. In Section 3, we address the equity premium puzzle, the risk-free rate puzzle and the excess-volatility puzzle, and present a simple calibration. We discuss robustness and how our approach is related to other approaches in Section 4. After a brief conclusion, all proofs appear in the Appendix, as does some supporting Mathematica code, which provides numerical back-up for our theoretical results.

2 Model

Consider an economy that evolves between times 0 and T, in which there are two sources of the consumption good. As in the standard one-tree model, the first, risky asset grows stochastically, and pays an instantaneous dividend of $D_t\,dt$, where $D_t = D_0\, e^{y(t)}$, $y(0) = 0$, $dy = \mu\, dt + \sigma\, d\omega$, and μ and σ are constants. Here, ω is a standard Brownian motion, which generates a standard filtration, $\mathcal{F}_t$, on $t \in [0, T)$. Unlike the one-tree model, there is also a second, riskless asset paying a dividend, $B\,dt$, where $B \geq 0$. It will be useful to consider the share of the risky asset in the overall economy and so we define the risky share, $z(t) = \frac{D_t}{B+D_t}$. We also define $\hat{\mu} = \mu + \frac{\sigma^2}{2}$. The horizon T can be finite or infinite. We focus primarily on the case when $T = \infty$, but show in Section 4 that the results carry over to the case with large but finite T. In Section 4, we also show how these assumptions on the growth processes can be substantially relaxed.

There is a price-taking representative investor with constant relative risk-averse (CRRA) utility, risk-aversion coefficient $\gamma > 1$, and personal discount rate $\rho > 0$, who consumes the total output:

$$U(t) = E_t \left[\int_t^{\infty} e^{-\rho(s-t)} u(B + D_s)\, ds \right], \tag{1}$$

where

$$u(c) = \frac{c^{1-\gamma}}{1-\gamma}. \tag{2}$$

We also write $U(t|B, D_t)$, when we want to stress the dependence on B and D_t. We note that when $\mu > 0$ and $B > 0$, the distribution of the risky share, $z(t) \in (0, 1)$, converges in probability to one for large t, $z \to_p 1$. In this case, the growth rate of real variables (i.e., dividends and consumption) in the economy behaves much like that in the one-tree model for large t. If, on the other hand, $\mu < 0$, the share converges to zero, $z \to_p 0$. In this case, real variables become almost risk free over time. If $\mu = 0$, then the share converges in probability to a two-point distribution with 50% mass at 0 and 50% mass at 1 (the convergence also holds a.s. for $\mu \neq 0$, but not for $\mu = 0$). In what follows, we focus our attention on the economically interesting case $\mu > 0$.

The market is dynamically complete, and usual arguments imply that, in equilibrium, an asset that pays out ξ_t, where ξ_t is an $\mathcal{F}_t$ adapted process satisfying standard conditions, commands an initial price of

$$P_0 = \frac{1}{u'(B + D_0)} E_0 \left[\int_0^\infty e^{-\rho s} u'(B + D_s) \xi_s \, ds \right]. \tag{3}$$

Equation (3) is the Euler equation relating the agent's aggregate consumption, marginal utility and valuation for all securities.

Notice that if $B = 0$, all resources are in the risky asset and the economy collapses to the standard one-tree model with constant growth and power utility. When $B > 0$, the economy is a special case of that in Cochrane, Longstaff, and Santa-Clara (2008), further generalized in Martin (2008); i.e., it is a so-called "two-trees" economy, in which one of the trees is risk free. We refer to the case $B = 0$ as the *standard model*, whereas when $B > 0$ we have the *Minimum Consumption* (MC) model.[5] As we elaborate below, providing the agent a minimal level of insurance (through the risk-free tree) provides new implications.

We define

$$\eta = \rho + (\gamma - 1)\mu - (\gamma - 1)^2 \frac{\sigma^2}{2},$$

the dividend yield in the standard model, which will be useful going forward. The properties of the standard model have been extensively analyzed, and are summarized in Table 1.

2.1 The Breakpoint

In the MC model, utility and marginal utility are bounded both from above and from below, so (1) and (3) are well-defined for arbitrary values of $\mu > 0$, $\sigma > 0$, $\rho > 0$, $B > 0$, and $D_0 > 0$. To clarify the differences between the MC model and the standard one, we study the expected utility of the agent in the two settings. First, observe that the homogeneity of the utility function implies that the value function, U, is scalable as $U(t|B, D_t) = (B + D_t)^{1-\gamma} U(t|1 - z, z) \stackrel{\text{def}}{=} (B + D_t)^{1-\gamma} w(z)$, where $w(z) \stackrel{\text{def}}{=} U(t|1 - z, z)$. We call $w(z)$ the *normalized* value function at share z.

[5] Equivalently, we could have specified the economy as one with HARA utility, $u(c) = \frac{(B+c)^{1-\gamma}}{1-\gamma}$, or one in which there is one asset with output $B + D_t$ and a risk-free bond in zero net supply.

Variable	Value
Risk-free rate, r_s	$\rho+\gamma\left(\mu+\frac{\sigma^2}{2}\right)-\gamma(\gamma+1)\frac{\sigma^2}{2}$
Long rate, r_l	$\rho+\gamma\left(\mu+\frac{\sigma^2}{2}\right)-\gamma(\gamma+1)\frac{\sigma^2}{2}$
Market return, r_e	$\rho+\gamma\left(\mu+\frac{\sigma^2}{2}\right)-\gamma(\gamma-1)\frac{\sigma^2}{2}$
Dividend yield, $\eta \stackrel{\text{def}}{=} D/P$	$\rho+(\gamma-1)\mu-(\gamma-1)^2\frac{\sigma^2}{2}$
Market risk premium, $r_e - r_s$	$\gamma\sigma^2$
Consumption volatility	σ
Dividend volatility	σ
Price volatility	σ
Market Sharpe ratio	$\gamma\sigma$

Table 1: Properties of the standard model (the consumption model with Brownian log-consumption process and power preferences).

We define the following three variables, which will be helpful going forward:

$$q=\sqrt{\mu^2+2\rho\sigma^2}, \qquad \kappa=\frac{\mu+q}{\sigma^2}, \qquad \alpha=\gamma-\kappa. \tag{4}$$

We shall see later that the value of α will be extremely important for the behavior of the model. Note that it is always the case that $\alpha<\gamma$.

Our first result characterizes the normalized value function.

Proposition 1 *In the MC model, the normalized value function of the representative agent, $w(z)$, is given by*

$$w(z) = \frac{z^{-\kappa}(1-z)^{1-\gamma-\kappa}}{q(1-\gamma)}\left[V\left(\frac{1-z}{z},\kappa,2-\gamma\right)\right. \tag{5}$$

$$\left. + \left(\frac{1-z}{z}\right)^{\frac{2q}{\sigma^2}} V\left(\frac{z}{1-z},\alpha+\frac{2q}{\sigma^2}-1,2-\gamma\right)\right]. \tag{6}$$

Here,

$$V(y,a,b) \stackrel{\text{def}}{=} \int_0^y t^{a-1}(1+t)^{b-1}\,dt \tag{7}$$

is defined for $a>0$. Also, $w(0)=\frac{1}{\rho(1-\gamma)}$.

Moreover, recall that the dividend yield, $\eta=\rho+(\gamma-1)\mu-(\gamma-1)^2\frac{\sigma^2}{2}$. Then, if $\eta>0$, $w(1)=\frac{1}{\eta(1-\gamma)}$. If, on the other hand, $\eta\le 0$, then $w(1)=-\infty$.

The function V is related to the incomplete Beta function, $B(x,a,b)\stackrel{\text{def}}{=}\int_0^x t^{a-1}(1-t)^{b-1}\,dt$ (see Gradshteyn and Ryzhik (2000)), via the relation $V(x,a,b)=(-1)^a B(-x,a,b)$. However, the Beta function is complex valued for negative values, so we prefer using the real-valued function V. Also, since the Beta function and the hypergeometric function satisfy the relationship $B(x, a, b) = {}_2F_1(1-b, a, a+1, x)$, we could equivalently have expressed the formula in terms of hypergeometric functions.

The last part of Proposition 1 is important. When $\eta > 0$, the value function in the MC model converges to that in the standard model as z approaches one. However, when $\eta \leq 0$, the two models behave completely differently. In this case, the value function is negative infinity in the standard model, and equilibrium is undefined. In contrast, the value function is always finite in the MC model. It is easy to check that the *breakpoint* at which the standard model becomes undefined occurs at the risk-aversion coefficient

$$\gamma = 1 + \kappa, \tag{8}$$

where κ is defined in (4). It is straightforward to check that $\gamma-(1+\kappa) > 0 \Leftrightarrow \alpha > 1 \Leftrightarrow \eta < 0$, so above the breakpoint the dividend yield in the standard model is formally negative.[6] Below the breakpoint point (i.e., for lower γ), the standard model is well-defined, whereas above the breakpoint it is not. Thus, although we may expect the MC model to converge to the standard model below the breakpoint, the characteristics of the MC model above the breakpoint are unclear.

To get a sense of what the breakpoint implies for risk aversion, suppose that the consumption growth rate, volatility of growth and personal discount rate are

$$\hat{\mu} = 0.75\%, \qquad \sigma = 4\% \qquad \text{and} \qquad \rho = 1\% \tag{9}$$

respectively.[7] With these parameters, Equation (8) shows that the breakpoint occurs at $\gamma = 10.6$, a not unreasonably high number.

3 Anomalies Revisited

Without loss of generality, we assume that $B \equiv 1$, i.e., that the risk-free part of the consumption stream is of size one, and from (3) we define $P(D_0)$ to be the price of the total

[6]It is well-known that expected utility is infinite beyond the breakpoint in the standard model. For example, Campbell (1986) develops a parameter restriction for general stationary processes that is the discrete time version of the breakpoint equation. The breakpoint condition also occurs in Martin (2008), though in a different context. Martin (2008) characterizes the prices of "small firms" below the breakpoint. We examine the properties of the market above the breakpoint.

[7]These numbers are quite close to international data for developed countries. For example, the average annual consumption volatility for ten countries between 1970–2000 reported in Campbell (2003) is 2.13%. Parker (2001) and Gabaix and Laibson (2001) argue that consumption adjustment costs may artificially reduce measured consumption volatility, and Mehra and Prescott (1985) use a volatility of 3.6%.

consumption output in the economy,

$$P(D_0) = E\left[\int_0^\infty e^{-\rho s}\left(\frac{1+D_0}{1+D_t}\right)^\gamma (1+D_t)\,ds\right]. \tag{10}$$

The price for general $B \neq 1$ then follows from the relation $P(B, D_0) = BP\left(\frac{D_0}{B}\right)$.

We provide an explicit characterization of the price of the market:

Proposition 2 *The price function $P(D)$ is*

$$P(D) = (1+D)^\gamma \frac{D^{-\kappa}}{q}\left[V(D,\kappa,2-\gamma) + D^{\frac{2q}{\sigma^2}} V\left(\frac{1}{D}, \alpha + \frac{2q}{\sigma^2} - 1, 2-\gamma\right)\right] \tag{11}$$

where $V(y,a,b) \stackrel{\text{def}}{=} \int_0^y t^{a-1}(1+t)^{b-1}\,dt$, and q, κ and α are defined as in Equation (4), i.e., $q = \sqrt{\mu^2 + 2\rho\sigma^2}$, $\kappa = \frac{\mu+q}{\sigma^2}$, $\alpha = \gamma - \kappa$.

Similar formulas are derived in Cochrane, Longstaff, and Santa-Clara (2008) (for $\gamma = 1$), and in Martin (2008), though there they are expressed in terms of hypergeometric functions.

Figure 1 shows the price-dividend ratio multiplied by $|\eta|$, for different choices of γ. This is one in the standard model, regardless of D. Recall that the breakpoint for this set of parameters is $\gamma = 10.6$. For $\gamma = 2$ and $\gamma = 3$ (the lower lines), the ratios quickly converge to

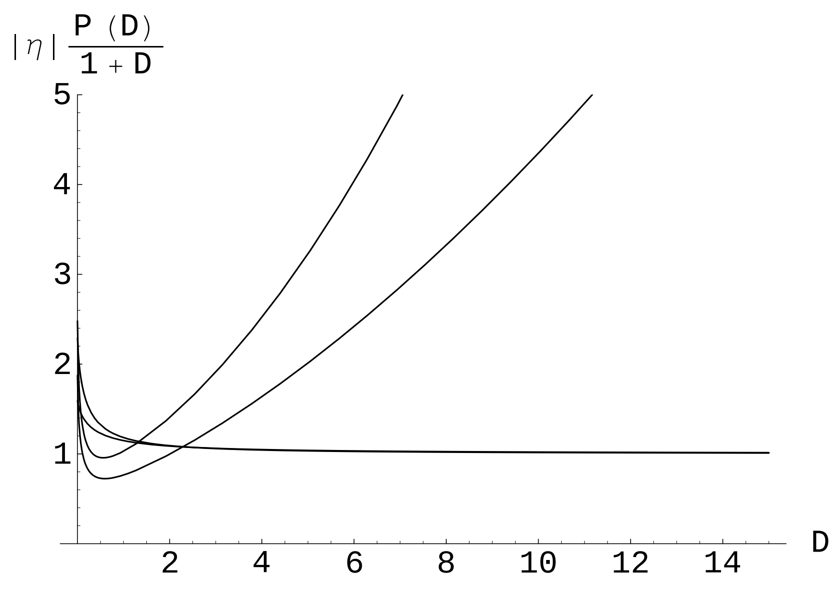

Figure 1: **Scaled market price-dividend ratio in MC model as a function of D, for parameters according to (9), i.e.,** $\hat{\mu} = 0.75\%$, $\sigma = 4\%\, \rho = 1\%$, **with risk aversion coefficient,** $\gamma = 2, 3, 12, 13$.

one as D increases, in line with the intuition that when D is large, the economy is effectively the same as the standard model. However, for $\gamma = 12$ and $\gamma = 13$, the function quickly increases as D grows. It is clear from the figure that price dynamics above the breakpoint are quite different from those below. We now explore why.

The existence of the risk-free tree clearly has a large impact on the pricing kernel in the relevant parameter range. Consider the price of a knock-in claim that pays a very small amount (say \$1) in the event that total consumption drops to $B + \epsilon$. Call the value of such an asset, which obviously depends on the starting point D_0, $K(D_0, \epsilon)$ (where we assume that $D_0 > \epsilon$).

It follows from (2) that K is given by

$$K(D_0, \epsilon) = \left(\frac{1 + D_0}{1 + \epsilon}\right)^{\gamma} E_0\left[e^{-\rho\tau_f}\right],$$

where τ_f is the stopping time

$$\tau_f \stackrel{\text{def}}{=} \inf_t \{t : D_t \le \epsilon\}. \tag{12}$$

The value of this claim is thus made up of two offsetting elements. The first element, $\left(\frac{1+D_0}{1+\epsilon}\right)^{\gamma}$, is the incremental marginal utility of the agent when he consumes, given his consumption today. The second element, $E_0\left[e^{-\rho\tau_f}\right]$, is the time value of \$1 when this state occurs. The first element increases with D_0, since the relative value of hitting ϵ is higher the wealthier the economy is at the starting point. The second element decreases with D_0, since the higher D_0, the longer it will take to reach ϵ (and the lower the probability that ϵ will ever be reached). How this valuation changes with changes in the underlying parameters is crucial to understanding why prices above the breakpoint are so different from those below.

It is straightforward, using standard results for stopping time distributions (see Ingersoll (1987)), to show that

$$K(D_0, \epsilon) = \left(\frac{1 + 1/D_0}{1 + \epsilon}\right)^{\gamma} \frac{D_0^{\alpha}}{\log[D_0]}, \tag{13}$$

where α was defined in (4). The first term is close to 1 for large D_0 and small ϵ, whereas the size of the second term depends on the sign of α. Below the breakpoint (i.e., for $\alpha < 1$), for large D_0 this asset is worth much less than D_0, i.e., $K(D_0, \epsilon)/D_0$ goes to zero as D_0 goes to infinity. In this case, the time-value of money effect dominates the marginal utility effect. Above the breakpoint, on the other hand (i.e., for $\alpha > 1$), this asset becomes very valuable for high D_0, in a nonlinear fashion. The marginal-utility effect now dominates the time-value-of-money effect.

The central intuition of the paper is that the trees contain this type of payout (they pay something in the bad states of the world). Therefore, above the breakpoint, the market value of these trees will also increase superlinearly with D_0.

The knock-in claim argument also provides an intuition for why the standard model does not work above the breakpoint. The single tree in the standard model also contains a collection of these types of threshold payments. The single tree does not, however, guarantee the representative agent the subsistence level of $B = 1$. The first term in the equation corresponding to (13) therefore only contains ϵ in the denominator. It follows that such claims will be much more valuable in the one-tree economy because when the agent's consumption is low (ϵ low), her marginal utility will be very high and therefore the value of such claims will explode. In this way, the risky tree becomes infinitely valuable. We will return to this point in more detail in Section 4, where we show that a similar argument also holds for finite horizon economies.

It is possible to derive the following asymptotic results for large D_0 of the behavior of the market price-dividend ratio in the MC economy.[8]

Proposition 3 *The asymptotic price-dividend ratio in the MC model depends on the parameter region. Specifically,*

(i) Below the breakpoint (i.e., for $\alpha < 1$ so that the value function is finite in the one-tree model), for large D the price-dividend ratio converges to $\frac{P(D)}{1+D} = \frac{1}{\eta}$.

(ii) Above the breakpoint (i.e., for $\alpha > 1$ so that the value function is infinitely negative in the one tree model), for large D the price-dividend ratio converges to $c\left(\frac{D}{B}\right)^{\alpha-1}$, for some constant $c > 0$, where α is defined in (4).

Proposition 3 shows that the the exponent of the asymptotic price-dividend ratio behaves like $\max(\alpha, 1) - 1$. It is thus the "convexity parameter," $\max(\alpha, 1)$, that governs the behavior of price-dividend ratios (and prices) for large D. Figure 2 below shows the convexity parameter as a function of risk aversion (γ) for some different parameter choices. The convexity of the price function lies at the heart of our analysis of the asset pricing anomalies, to which we now turn.[9]

3.1 The risk premium

It is important to stress that reasonable values of the exogenous parameters are consistent with the region in which prices and price-dividend ratios are undefined in the standard

[8]Throughout the paper we study the value of the total $B + D$ consumption flows. We obtain identical asymptotic results for the purely risky part of the economy, i.e., the value of the D consumption flows.

[9]The convexity of the price function above the breakpoint (shown in Proposition 3(ii)) is crucial for the subsequent results. The convexity can also be verified numerically. We provide Mathematica code for the numerical calculations in the Appendix.

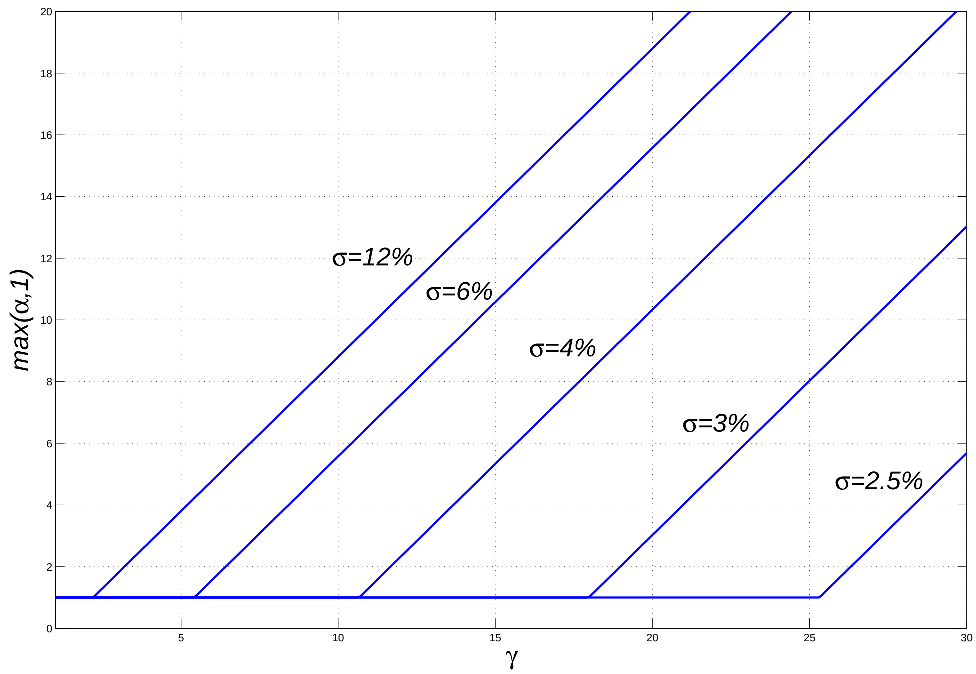

Figure 2: **Convexity parameter,** $\max(\alpha, 1)$**, as a function of risk aversion** γ**. Parameters:** $\hat{\mu} = 0.75\%$**,** $\rho = 1\%$**,** $\sigma = 2.5\%, 3\%, 4\%, 6\%, 12\%$**.**

model. In the MC economy, the asymptotic expected return on the market depends on the parameter values. Recall that the instantaneous expected return on the market is

$$r_e\, dt \quad = \quad E\left[\frac{dP}{P} + \frac{1+D}{P}\, dt\right]. \tag{14}$$

We have

Proposition 4 *For large D*

(i) Below the breakpoint, the expected return on the market is the same as in the standard model: $r_e = \rho + \gamma\mu - \gamma(\gamma - 2)\frac{\sigma^2}{2}$.

(ii) Above the breakpoint, the expected return on the market is $r_e = \alpha\mu + \alpha^2\frac{\sigma^2}{2}$, *where* α *is defined in (4).*

To get an intuition for the results in Proposition 4, we note that below the breakpoint, the price is essentially the same as in the standard model (as shown in Proposition 3(i)), so expected returns will essentially be the same. Above the breakpoint, however, the second term in (14) becomes small for large D (as implied by Proposition 3(ii)). Moreover, since $P(D) \sim D^\alpha$, the first term behaves like

$$E\left[\frac{dP}{P}\right] = \frac{(\mu + \frac{\sigma^2}{2})P'\,dt + \frac{\sigma^2}{2}D^2P''\,dt}{P} \approx \alpha(\mu + \frac{\sigma^2}{2})\,dt + \alpha(\alpha-1)\frac{\sigma^2}{2}\,dt.$$

It follows that the market risk premium can also become large. In fact, it is well known that the risk-premium, $r_e - r_s$ can be expressed as

$$(r_e - r_s)\,dt = -cov\left(\frac{dM}{M}, \frac{dP}{P}\right),$$

where r_s is the short-term rate, and M is the pricing kernel, which is equal to $e^{-\rho t}(1+D)^{-\gamma}$ in the MC economy. It therefore follows from standard Itô calculus that

$$r_e - r_s = \gamma \max(\alpha, 1)\sigma^2. \tag{15}$$

For $\alpha < 1$, the risk premium is thus the same as in the standard model. For $\alpha > 1$, however, it is larger. In this case, through α, the risk premium now depends on the economy's growth rate, μ, and the personal discount factor, ρ, and is decreasing in both of these parameters. Moreover, for low values of μ and ρ, the risk premium approaches $\gamma^2\sigma^2$ instead of $\gamma\sigma^2$.

These are asymptotic results, for large D. In Figure 3 we illustrate the market risk premium for a fixed risk aversion, $\gamma = 12.25$ (which is above the breakpoint), as we vary the risky share, z (recall that $z = \frac{D}{D+B} \in (0,1)$), using the parameters in (9). As D becomes large there is indeed convergence to the asymptotic value of 5.0%. Comparing this with the risk premium implied by the standard model, $\gamma\sigma^2 = 2.0\%$, we see that the premium in the MC model is substantially higher.

The following Figure 4 displays the market risk premium for a large D as we vary both risk aversion and volatility (each curve corresponds to a different volatility). Beyond the breakpoint, the risk premium increases very quickly in a convex fashion, implying that a small increase in risk aversion drastically increases the market risk premium.

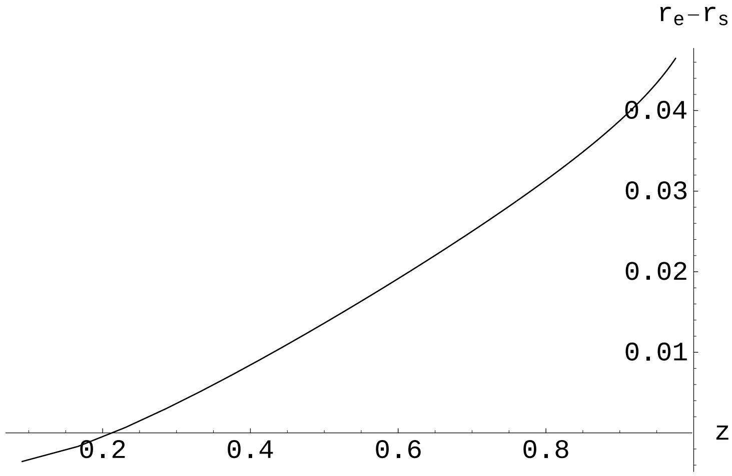

Figure 3: **Market risk-premium as a function of z, for fixed risk aversion. Parameters:** $\hat{\mu} = 0.75\%$, $\rho = 1\%$, $\sigma = 4\%$, $\gamma = 12.25$, **implying that** $\alpha = 2.55$. **The asymptotic risk-premium is** $\gamma\alpha\sigma^2 = 5.0\%$.

3.2 The term structure

The term structure is also quite different in the MC economy. From (3) it follows that a zero-coupon risk-free bond with maturity date τ has the price

$$P^\tau = e^{-\rho\tau} E_0 \left[\left(\frac{B + D_0}{B + D_t} \right)^\gamma \right]. \tag{16}$$

We can rewrite this expectation in terms of the risky share, $z = \frac{D_0}{B+D_0}$,

$$P^\tau = e^{-\rho\tau} E_0 \left[\left(1 + z \left(\frac{D_t}{D_0} - 1 \right) \right)^{-\gamma} \right], \tag{17}$$

and since the distribution of $\frac{D_t}{D_0}$ does not depend on D_0, it immediately follows that the price can be written as a function of z alone, $P^\tau(z)$. The following Proposition provides a price formula:

Proposition 5 *Define the log-relative size of the sectors* $d = \log(z/(1-z))$. *Then the price of a* τ*-period zero-coupon bond is given by*

$$P^\tau = (1 + e^d)^\gamma e^{-\rho\tau} \frac{1}{\sqrt{2\pi\sigma^2\tau}} \int_{-\infty}^{\infty} \frac{e^{-(y-d-\mu\tau)^2/(2\sigma^2\tau)}}{(1+e^y)^\gamma} dy. \tag{18}$$

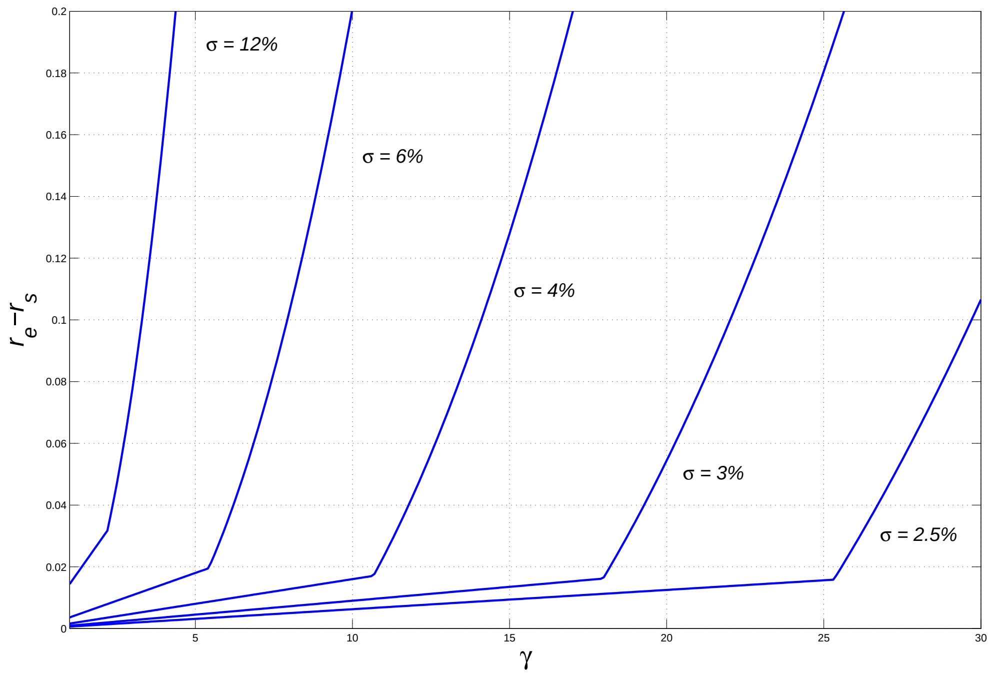

Figure 4: **Market risk premium for large D, as a function of risk aversion, γ. Below the breakpoint, the risk premium is the same as in the standard model, and linear in γ, $r_e - r_s = \gamma\sigma^2$. Above the breakpoint, the risk premium is a steeply convex function. Parameters: $\hat{\mu} = 0.75\%$, $\rho = 1\%$, $\sigma = 2.5\%, 3\%, 4\%, 6\%, 12\%$.**

By defining $F(x) = e^{x^2} Erfc(x)$, where Erfc is the error function, $Erfc(x) = (\sqrt{\pi})^{-1} \int_z^\infty e^{-t^2}\, dt$, Equation (18) can be expressed in the following form:

$$P^\tau = \frac{(1+e^d)^\gamma e^{-\rho\tau - (d+\mu\tau)^2/(2\sigma^2\tau)}}{2} \times \lim_{\epsilon \searrow 0} \sum_{n=0}^{\infty} (-1)^n e^{-\epsilon n} a_n \left(F\left(\frac{\epsilon + d + \mu\tau + n\tau\sigma^2}{\sqrt{2\sigma^2\tau}} \right) + F\left(\frac{\epsilon - d - \mu\tau + (n+\gamma)\tau\sigma^2}{\sqrt{2\sigma^2\tau}} \right) \right), \quad (19)$$

Here,

$$a_n = \frac{\Gamma(\gamma+n)}{\Gamma(\gamma)\Gamma(n+1)},$$

where $\Gamma(x) = \int_0^\infty e^{-t} t^{x-1}\, dt$, which reduces to $a_n = \binom{n+\gamma-1}{\gamma}$ when γ is integer valued.

Martin (2008) independently characterizes the term structure in an economy with many trees. His framework is more general than ours, in that it allows for general Levy processes and multiple trees, but his solution is based on Fourier transform techniques, and so is different from those in Proposition 5.

We derive the long and short rates from the price of the zero-coupon, risk-free bond that we established in Proposition 5. Specifically, the τ-period spot rate is defined as

$$r_\tau = -\frac{\log(P^\tau)}{\tau},$$

while the short and long rates are defined to be

$$r^s = \lim_{\tau \searrow 0} r_\tau, \qquad \text{and} \qquad r^l = \lim_{\tau \to \infty} r_\tau,$$

respectively.

In the MC economy, the term structure is no longer constant. Indeed, it can slope upwards or downwards, and can even be hump-shaped. We use Equation (19) to study the yield curve with parameters chosen according to (9), $z = 70\%$, and risk-aversion coefficients between 6 and 12. The results are shown in Figure 5. The choice of $z = 70\%$ means that

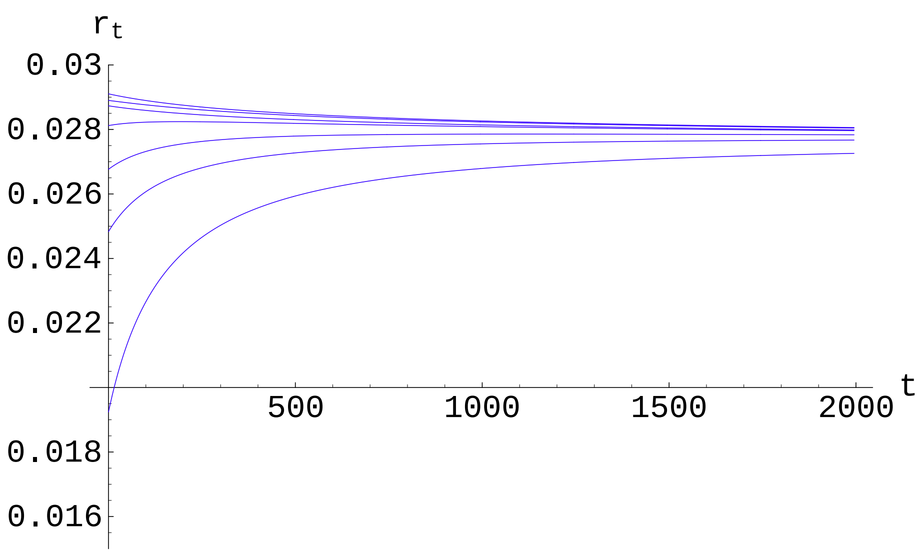

Figure 5: **Term structure of interest rates in the MC model. Parameters, $z = 0.7$, γ varies between 6 (highest curve) to 12 (lowest curve). Other parameters according to (9).**

the risky tree initially dominates the economy, and the risky share converges to $z = 1$ as t grows, so the consumption growth rate is fairly stable in this economy.

We note that the yield curves in the figure are upward sloping and that the slope and curvature increase with the risk-aversion coefficient, γ. Moreover, although the short end of the curve is sensitive to γ, as in the one-tree model, there seems to be an asymptotic long-term rate that does not seem to vary much with γ. To understand these properties of the yield curve, we analyze the short rate, r^s, and the long rate, r^l. We have

Proposition 6 *In the MC economy, the short-term rate is*

$$r^s = \rho + \gamma z \left(\mu + \frac{\sigma^2}{2} \right) - \gamma(\gamma + 1)\frac{\sigma^2}{2} z^2.$$

For $z \in (0, 1)$, if $\mu \leq \gamma\sigma^2$, the long-term rate is

$$r^l = \rho + \frac{1}{2} \times \frac{\mu^2}{\sigma^2}. \tag{20}$$

If, on the other hand, $\mu > \gamma\sigma^2$, the long-term rate is

$$r^l = \rho + \gamma \left(\mu + \frac{\sigma^2}{2} \right) - \gamma(\gamma + 1)\frac{\sigma^2}{2}. \tag{21}$$

Thus, the short rate has the same structure as in the standard model and, as long as $\mu > \gamma\sigma^2$, the long rate is the same as in the standard model. This makes intuitive sense, since the economy will almost surely be very similar to the one-tree economy in the long run. If $\mu < \gamma\sigma^2$, however, the long rate is a constant, independent of the risk aversion parameter. Since

$$\eta = \rho + (\gamma - 1)\mu - (\gamma - 1)^2\frac{\sigma^2}{2} > (\gamma - 1)\left(\mu - (\gamma - 1)\frac{\sigma^2}{2} \right) > (\gamma - 1)\left(\mu - \gamma\sigma^2 \right),$$

it will always be the case that the long rate is independent of risk-aversion above the breakpoint (i.e., when η is negative).

In our previous numerical example, with parameters according to (9) and $\gamma = 12.25$, this implies that the long rate is $r_l = 2.4\%$. The short-rate depends on z, as shown in Figure 6. For z close to unity, i.e., for large D, it becomes negative. At $z = 1$, it is -2.8%. While negative, this is nevertheless far more reasonable than the values we would obtain if we calibrated the standard model to the market risk premium. For example, a risk premium of 5% would imply a risk-free rate of -58% in the standard model. We are, of course, dealing with real variables, so a negative discount rate is obviously possible, although not what we see in practice. For $z \leq 0.8$, the short rate is positive though. From Figure 3, we see that the risk-premium is about 3% at $z = 0.8$. Finally, we note that the volatility of the short rate, $\sigma(r_s)$, is low and depends on z. In our example, $\sigma(r_s)$ varies between 0 and 0.08% and reaches its maximum at $z \approx 0.75$.

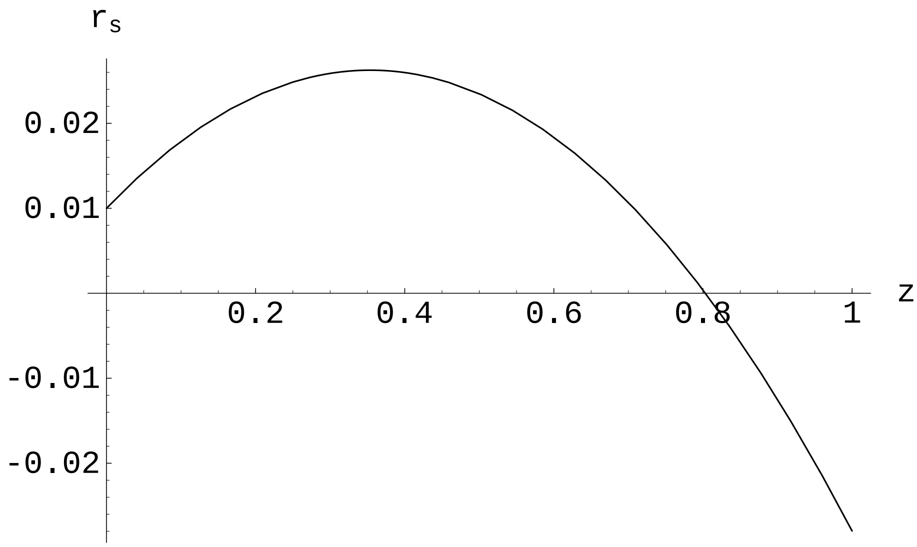

Figure 6: **Short rate as a function of z. Parameters according to (9) and $\gamma = 12.25$.**

This γ-independence above the breakpoint stands in stark contrast to the results in the standard model, where the interest rate is very sensitive to risk aversion. Specifically, in the MC model, the long rate is always greater than the personal discount rate, $r_l > \rho$, regardless of the aggregate risk aversion in the economy, and is therefore positive.[10,11] This γ-independence thus offers a resolution to the risk-free rate puzzle at the long end of the term structure.

The reason why risk aversion becomes unimportant for bond yields as the horizon increases, even though bond prices depend on risk aversion, is that differences between bond prices in economies indexed by different levels of risk aversion are sufficiently small, compared with the compounding inherent in the yield calculation, that the price differences become unimportant at the long end of the curve. The price of a bond is the expected discounted value of a dollar multiplied by the representative agent's marginal utility. In

[10]A somewhat related result on the long rate is presented in Dybvig, Ingersoll, and Ross (1996), who show that long rates can never fall over time because Bayesian updaters can never be surprised by a worse state. Within our specific economy, our result is stronger than the Dybvig-Ingersoll-Ross theorem, since it states that r^l is constant over time, and across risk-aversion.

[11]We have verified that the formula is indeed correct by numerically integrating Equation (16) directly. Mathematica code is provided in the appendix, showing that with parameters, $\rho = 1\%$, $\mu = 3.5\%$, $\sigma = 20\%$, $\gamma = 2.5$, the long rate converges to $r^l = \rho + \frac{\mu^2}{2\sigma^2} = 2.53\%$ (in line with Equation (20), since $3.5\% < 2.5 \times 20\%^2$). On the contrary, Equation (21) would, for example give $r^l = \rho + \gamma\mu - \gamma^2\sigma^2/2 = 1\% + 2.5 \times 3.5\% - 2.5^2 \times 20\%^2/2 = -2.75\%$. By varying B, D_0 and γ, it is easily verified that r^l does not depend on these parameters.

the MC model, the marginal utility (irrespective of risk aversion) is bounded above and below. If the agent consumes the fruit of a risk-free tree, which provides insurance, then marginal utility is always bounded above. Indeed, one can find an upper bound on the ratio of marginal utilities for two agents with the same personal discount factor but different risk aversion coefficients independently of time horizon. Therefore, bond prices for the same maturity for any two economies that differ only in the risk aversion of their representative agents will not differ "by much." For long maturities this will lead to similar yields.

The difference between the long rates in the standard and MC economy further underlines the fragility of the CRRA-lognormal model in longer time horizons. Regardless of how close z is to 1 in the MC model, the long-term rate is drastically different from when z is identically equal to 1. The differences between the two models are driven by the insurance the risk-free tree provides in the far-left tails. Moreover, although the long rate is always γ-independent above the breakpoint, there are also regions *below* the breakpoint in which it is γ-independent.

At a broad level, our results are reminiscent of, but distinct from, those found in Weitzman (1998, 2001). Weitzman argues that if there is parameter uncertainty, the long-term discount rate is lower than that inferred from the short- and mid-term rates We agree with Weitzman that a careful analysis of the implicit assumptions about return distributions and utility in the tails is needed to understand the long-term discount rate. Both Weitzman's and our results are driven by the extreme importance of the worst states in longer horizons. Unlike in Weitzman (1998, 2001), however, the long rate in our model may be higher than the short rate. This distinction is obviously important if existing market data are used to infer a maximum possible discount rate.

3.3 Excess Volatility

Above the breakpoint, prices are not linearly related to consumption, but, as we observed in Proposition 2, are a convex function of dividends. It naturally follows, then, that the volatility of prices is much higher than the volatility of the underlying dividends. In fact, it is easy to show that the price volatility is

$$vol\left(\frac{dP}{P}\right) = \max(\alpha, 1)\sigma. \tag{22}$$

In our numerical example, with parameters according to (9) and $\gamma = 12.25$, this implies a market price volatility of 10.3%, which is more than 2.5 times the dividend (and consumption) volatility of 4%. The model thus naturally leads to excess volatility, both with respect to consumption and with respect to dividends. Since $\alpha < \gamma$, an upper bound on the excess volatility is given by the risk-aversion parameter.

In Table 2, we summarize the formulas and numerical results we have derived, and also include the corresponding results from the standard model (see Table 1).

Variable	**Formula**	**Value-MC**	**Value-standard**[a]
Short rate, r_s	$\rho + \gamma\left(\mu + \frac{\sigma^2}{2}\right) - \gamma(\gamma+1)\frac{\sigma^2}{2}$	-2.8%	-2.8%
Long rate, r_l,when $\mu < \gamma\frac{\sigma^2}{2}$	$\rho + \frac{\mu^2}{2\sigma^2}$	2.4%	-2.8%
Long rate, r_l, when $\mu > \gamma\frac{\sigma^2}{2}$	$\rho + \gamma\left(\mu + \frac{\sigma^2}{2}\right) - \gamma(\gamma+1)\frac{\sigma^2}{2}$		
Market return, r_e, when $\alpha > 1$	$\alpha\mu + \alpha^2\frac{\sigma^2}{2}$	2.2%	-0.8%
Market return, r_e, when $\alpha < 1$	$\rho + \gamma\left(\mu + \frac{\sigma^2}{2}\right) - \gamma(\gamma-1)\frac{\sigma^2}{2}$		
Risk premium, $r_e - r_s$	$\gamma \max(\alpha, 1)\sigma^2$	5.0%	2.0%
Consumption volatility	σ	4%	4%
Dividend volatility	σ	4%	4%
Price volatility	$\max(\alpha, 1)\sigma$	10.3%	4%
Market Sharpe ratio	$\gamma\sigma$	0.49	0.49

Table 2: Properties of the MC model for large D, and an example with parameters according to (9), $\hat{\mu} = 0.75\%$, $\sigma = 4\%$, $\rho = 1\%$, with $\gamma = 12.25$, implying that $\alpha = 2.58$.

[a] From formulas given in Table 1. Beyond the breakpoint, the equilibrium is not defined in the standard model, although the values can still formally be calculated.

4 Robustness of Results

4.1 Sensitivity of standard model

Our model shows how changing the process in low-consumption states drastically changes the results. The importance of these states in the CRRA-lognormal framework was emphasized in Kogan, Ross, Wang, and Westerfield (2006), who studied the price impact of irrational traders in capital markets. Earlier, Geweke (2001) made the point that the CRRA-lognormal framework is not robust to different distributional assumptions in the far-left tails. In fact, a small "fattening" of the left-tail distribution, introduced, for example, by parameter uncertainty in a Bayesian framework, makes expected utility infinitely negative.

Further, Barro (2005) (following Rietz (1988)) uses the absence of consumption insurance to generate empirically reasonable risk premia. Along somewhat similar lines, Weitzman (2007) argues that parameter uncertainty can explain the equity premium puzzle. At first blush, we appear to be doing the opposite of Rietz (1988) and Barro (2005), who suggested that rare, catastrophic events could drive the risk premium, and also to Weitzman (2007), since parameter uncertainty makes the representative investor increase the probability weights on low-consumption outcomes. However, the intuition behind our resolution is, in spirit, similar to theirs. In both cases, a higher risk premium arises because very bad outcomes are more important than in the standard model. Unlike Rietz (1988), Barro

(2005) and Weitzman (2007), our modification actually *lowers* the probability of extremely bad outcomes, but this allows us to study regions of parameter space, invalid under the standard model, where bad events have a much larger effect on expected utility. In these regions the risk premium is higher even though there is no "jump risk" in the MC model.

4.2 Finite time horizons

The standard model is not defined above the breakpoint in the infinite-horizon setting. It is, however, well defined above the breakpoint when the time horizon is finite, with the same low market risk premium, $r_e - r_s = \gamma\sigma^2$, as below the breakpoint. Similarly, it is straightforward to show that the MC economy with long but finite horizon converges to the MC economy with infinite horizon.[12] How can the results then be so different?

We argue that it is the standard model that behaves strangely above the breakpoint. The price function in the finite-horizon case is $P(t, D) = \frac{1-e^{-\eta(T-t)}}{\eta} D$. Below the breakpoint, this converges to $\frac{D}{\eta}$ for large T. Above the breakpoint, on the other hand, the price explodes as time to maturity increases. The low risk premium then comes from the fact that $\frac{P_t}{P} \approx \eta$, i.e., there is a large expected price decrease at each point in time when η is negative. This decrease is driven by the low-state knock-in claims we discussed previously. These claims are extremely valuable for long time horizons, but their value decreases very quickly over time when the terminal date approaches, since the risk that these states will ever be reached decreases rapidly. The behavior of the entire tree's value is driven in large part by the extreme behavior of these low-state knock-in claims. Since such negative expected returns with time horizon do not seem to be present in practice, we conclude that the standard model also provides a poor characterization above the breakpoint with finite horizons.

4.3 Relation to literature on bubbles

Price-dividend ratios in the MC model are nonstationary beyond the breakpoint. In fact, the convex price function is similar for large D to what occurs in the rational-bubble models that have been introduced to explain the excess-volatility puzzle (see, for example, Froot and Obstfeld (1991)). In the MC economy, however, even though price-dividend ratios are nonstationary, there is no bubble, since the discounted cash flow formula (3) prices all assets in the economy. In fact, as noted already in Cochrane (1992), Appendix B, even with stationary distributions for consumption growth, price-dividend ratios need not be stationary.

The empirical literature that has tested for explosive stock market price dynamics has produced mixed results. For example, Diba and Grossman (1988) use a cointegration aug-

[12] The convergence follows much easier than in the standard model since, for $\gamma > 1$ and $B > 0$, the utility function and its derivative are bounded below and above for all states of the world.

mented Dickey-Fuller test to conclude that prices are not explosive, a conclusion that is supported by Cochrane (1992). On the other hand, West (1987) and Froot and Obstfeld (1991) do find evidence for explosive price dynamics, findings that are also supported by Engsted (2006), who uses a cointegrated VAR method. In the MC model, the price-dividend ratios explode quite slowly and may therefore be hard to detect. In our numerical example, for example, it takes about 65 years for price-dividend ratios to double. This compares with an observed increase in the market price-dividend ratio of 3.2 times during the 65 years between 1943 and 2008.[13]

4.4 Relation to Hansen-Jagannathan bounds

We have developed our results with respect to the market risk premium. In other words, our analysis has rested on the assumption that the equity portfolio makes up the whole market portfolio. This is the formulation developed in Mehra and Prescott (1985) and many other papers. With that formulation it is shown that, all else equal, the risk-premium is much higher in the MC model than what seems to be implied by the standard model.

An alternative approach to the equity premium is given in Hansen and Jagannathan (1991), in which it is described as a bound on the Sharpe ratio of the equity portfolio. This bound puts restrictions on the SDF in the economy, whereas the market model puts joint restrictions on the SDF and the price function. Since the risk premium in our approach increases due to a more volatile price function, our approach therefore has less to say about the Hansen-Jagannathan bounds. It does have two implications though. First, the interpretation of a high equity volatility differs from that in the standard model. In the standard model, a high equity volatility implies that the equity market is a highly leveraged claim on consumption. This is not the case in the MC model, in which the high volatility is introduced because of the convex price function. Second, since the price function is nonlinear, the unconditional correlation between consumption and equity returns may be low even though the two processes are instantaneously perfectly correlated. In fact, it follows from Figure 1 that for low D, equity and consumption are perfectly negatively correlated, which will decrease the unconditional correlation and, in turn, artificially make the required risk premium look higher than it actually is (see Berk and Walden (2009) for further analysis of this argument).

4.5 Generalizations

For simplicity, we have derived our results in a two-trees framework with one risk-free, constant-size tree. The results, however, are much more general. As long as there is a lower

[13] This calculation is based on annual price and dividend data obtained from Robert Shiller's Web site, http://www.econ.yale.edu/~shiller/data.htm.

bound on consumption (which could grow deterministically at some rate), and the risk of ending up in these low-consumption states is bounded below by an i.i.d. growth process that, given γ, is above the breakpoint, similar results apply. The general consumption process could, for example, contain mean-reverting growth, as well as long-term i.i.d. growth. To fix ideas, we illustrate one such generalization and show how a convex price function arises under general conditions.

Proposition 7 *Consider an exchange economy, with a representative agent with CRRA expected utility with risk aversion coefficient $\gamma > 1$ and personal discount rate $\rho > 0$, in which the consumption is $C_t = f(D_t)$, where $D_t = e^{s_t}$, and where $f : \mathbb{R}_+ \to \mathbb{R}_+$ is a continuous, increasing function, such that for large d, $c_0 d \leq f(d) \leq c_1 d$, for some constants $0 < c_0 \leq c_1 < \infty$.*

For the stochastic process, $s_t \in \mathbb{R}$, define the c.d.f. $F(s,t|s_*,\mathcal{I}) = \mathbb{P}(s_t \leq s | s_0 = s_*, \mathcal{I})$, where $\mathcal{I}$ captures the information known about s_t at $t = 0$. Assume that the following condition is satisfied:

$$\exists\mu, \exists\sigma > 0,\ \exists\bar{t} \geq 0,\ \exists\underline{s},\ \exists\overline{s},\ \forall t \geq \bar{t},\ \forall s_* \geq \overline{s},\ \forall s < \underline{s} : F(\underline{s}, t|s_*, \mathcal{I}) \geq \Phi\left(\frac{s - s_* - \mu t}{\sigma\sqrt{t}}\right). \quad (23)$$

Here, Φ is the cumulative normal distribution function. Further, assume that the economy is beyond the breakpoint, i.e., that $\alpha = \gamma - \frac{\mu + \sqrt{\mu^2 + 2\rho\sigma^2}}{\sigma^2} > 1$.

Then

(i) If $f(x) \leq c_2 x$ in a neighborhood of $x = 0$, for some constant $c_2 \geq 0$, then there is no equilibrium in the economy.

(ii) If $f(0) > 0$, then in any equilibrium the price of the market satisfies $P(C_0) \geq c_3 C_0^{\alpha}$, for some constant $c_3 > 0$.

Equation (23) states that for large D_0 (i.e., for $D_0 \geq e^{\overline{s}}$) and large t (i.e., for $t \geq \bar{t}$), the risk of ending up in low-consumption states ($s_t \leq \underline{s}$) is at least as high as if s were a constant coefficient Brownian motion with growth rate μ and volatility σ.

Example 1 *The MC economy is a special case of Proposition 7, in which $f(x) = 1 + x$, and $s_t \sim N(s_0 + \mu t, \sigma^2 t)$. It therefore satisfies (23) as an equality for all $\bar{t} > 0$, $\underline{s}$, and $\overline{s}$.*

Moreover,

Example 2 *Consider an MC economy with* $f(x) = 1 + x$ *and a mean-reverting growth process,*

$$
\begin{aligned}
ds_t &= \mu_t\, dt + \sigma\, d\omega_1, \\
d\mu_t &= \beta(\mu - \mu_t)\, dt + \sigma_\mu\, d\omega_2,
\end{aligned}
$$

where μ, σ, σ_μ, *and* β *are positive constants and where* $cov(d\omega_1, d\omega_2) = \rho\, dt$. *Similar processes are, for example, assumed in Kim and Omberg (2002), Wachter (2002), and Bansal and Yaron (2004), and it is well-known that the distribution of* s_t *is normally distributed,*

$$
\begin{aligned}
s_t \sim N\Bigg(& s_0 + \mu t + \frac{\mu_0 - \mu}{\beta}\left(1 - e^{-\beta t}\right), \frac{\sigma_\mu^2 + 2\sigma_\mu\beta\rho\sigma + \beta^2\sigma^2}{\beta^2} t \\
& - \frac{3\sigma_\mu^2 + 4\sigma_\mu\beta\rho\sigma}{2\beta^3} + \frac{2e^{-\beta t}}{\beta^3}\left(\sigma_\mu^2 + \sigma_\mu\beta\rho\sigma\right) - \frac{e^{-2\beta t}\sigma_\mu^2}{2\beta^3}\Bigg).
\end{aligned}
$$

Therefore, for large t, *(23) is satisfied with* $\sigma^2 \stackrel{\text{def}}{=} \frac{\sigma_\mu^2 + 2\sigma_\mu\beta\rho\sigma + \beta^2\sigma^2}{\beta^2}$. *Here,* $\sigma > 0$, *as long as* $\rho > -1$ *or* $\sigma_\mu \neq \beta\sigma$. *From Proposition 7 it therefore follows that similar price-dynamics occur beyond the breakpoint in the MC economy with a mean reverting growth process.*

Thus, our theory is really about minimum consumption levels in exchange economies, not about specific tree economies. In particular, referring back to the discussion after Equation (15), this result implies that if the representative investor has a low discount rate and believes that growth will slow down some time (arbitrarily far) into the future, then the effective equity premium for high D will still be approximately $\gamma^2\sigma^2$, regardless of the value of μ today (since it is only the asymptotic growth that matters). With this line of reasoning, the observed risk-premium in the example we have studied throughout this paper would be matched by $\gamma = \sqrt{5\%/4\%^2} = 5.6$ (instead of $\gamma = 12.25$ needed when the expected growth rate is constant). Further, if we use the numbers in Weitzman (2007) — a risk premium of 6% and consumption volatility of 2% — the risk premium is matched by $\gamma = \sqrt{6\%/2\%^2} = 12.25$ (compared with $\gamma = 6\%/2\%^2 = 150$ obtained in Weitzman (2007) under the assumption that $r_e - r_s = \gamma\sigma^2$). Thus, a high risk premium may be a sign that the economy will not be able to continue to grow fast in the long run.

5 Concluding Remarks

We have established that if risk aversion is sufficiently high, the stochastic discount factor in a simple one-tree exchange economy with minimum consumption can be a convex function of the dividend (and hence consumption) stream. This immediately leads to explosive price-dividend ratios, excess volatility, modest interest rates, and risk premia that are in line with those observed.

Intuitively, there are two main channels through which future low-consumption states affect how the representative agent values the market. The first is how the representative agent currently values these low states, and is therefore captured by the difference between marginal utility at current consumption and at the low-consumption states; the higher the current consumption, the greater the difference. Further, since marginal utilities are convex functions of consumption (when risk aversion is greater than one), this channel also makes current market prices convex in consumption. The second channel is how likely the representative agent is to hit one of these low states; the higher her current consumption, the lower the risk that the low-consumption states will ever be reached (and the longer it will take if they are reached). Below the so-called breakpoint, the second effect outweighs the first, which means that the influence of the consumption provided in low-consumption states on the current price becomes negligible when current consumption is high. This corresponds to the standard model, in which the value of the agent's consumption stream is essentially linear in that consumption. However, when risk-aversion is high enough to be above the breakpoint, the first effect dominates: The value of being able to consume in the low-consumption states *increases* convexly as current consumption grows. There is a ready analogy to this intuition in the rare-disaster literature; while there are no "disasters " in this framework, the existence of low-consumption states completely changes the properties of the model above the breakpoint.

There are two immediate conclusions that can be drawn from our work. First, the standard long-horizon one-tree model with a CRRA representative investor and a lognormal consumption process is highly sensitive to small perturbations, especially when risk aversion is high. In short, the framework is not robust. Second, an economically plausible assumption that is quite innocuous (subsistence consumption) renders predictions that are more in accord with empirical work. Of course, this one augmentation does not solve all puzzles; the short term risk-free rate is still too low, consumption volatility a bit too high, as is the coefficient of risk-aversion. However, we find it fascinating that such a small modification of the classic work-horse consumption model can improve the "fit" so significantly.

Finally, and more broadly, our results indicate that there is yet more to learn about the effect of the consumption process on asset prices. Because consumption (as opposed to utility) is observable; exhausting the implications of tractable models with plausible consumption streams presents a fruitful research agenda.

Proofs

Proof of Proposition 1:

The functional form for $z \in (0,1)$ follows from Proposition 2. Since, for $z = D_0/(B+D_0)$,

$$w(z) = (B+D_0)^{\gamma-1}U(0|B,D_0) = \frac{1}{1-\gamma}\frac{P(B,D)}{B+D} = \frac{1}{1-\gamma}\frac{B}{B+D}P\left(\frac{D_0}{B}\right) = \frac{1}{1-\gamma}zP\left(\frac{z}{1-z}\right),$$

where the second equality holds for for CRRA utility, which follows from (3). Therefore, (11) immediately leads to (6).

For the results at $z = 0$ and $z = 1$, we define $\hat{w}_T(z) = E\left[\int_0^T e^{-\rho t}u\left(1-z+ze^{y_t}\right)dt\right]$, where $y_t = \log(D_t/D_0)$. Thus, $w(z) = \hat{w}_\infty(z)$. It follows immediately that $w(1) = \hat{w}_\infty(1) = \int_0^\infty \frac{e^{-\rho t}}{1-\gamma}dt = \frac{1}{\rho(1-\gamma)}$. Moreover, $\hat{w}_T(0) = \int_0^T \frac{e^{-\eta t}}{1-\gamma}dt = \frac{1-e^{-\eta T}}{\eta(1-\gamma)}$, so for $\eta > 0$, $w(0) = \hat{w}_\infty(0) = \frac{1}{\eta(1-\gamma)}$, whereas for $\eta < 0$, $\lim_{T\to\infty}\hat{w}_T(0) = -\infty$. The proposition is proved.

We note that although $\hat{w}_\infty(0) = \lim_{T\to\infty}\lim_{z\to 0}\hat{w}_T(z) = -\infty$ when $\eta < 0$, it does not immediately follow that $\lim_{z\to 0} w(z) = \lim_{z\to 0}\lim_{T\to\infty}\hat{w}_T(z)$ is equal to $-\infty$ (for example, if $\hat{w}_T(z) = -\frac{1}{zT}$, then the former expression is infinite, whereas the second is zero). However, the latter result follows, since $\hat{w}_T(z)$ is decreasing in T for arbitrary $z \in [0,1]$, and $\hat{w}_T(z)$ is continuous in z for arbitrary finite T. Specifically, for an arbitrary constant, $k > 0$, it follows that for T^* large enough, $\hat{w}_{T^*}(0) \leq -2k$, and because of the continuity in z, $\hat{w}_{T^*}(z) \leq -k$ for all $z \leq z^*$, for some $z^* > 0$. Therefore, $\hat{w}_\infty(z) \leq \hat{w}_{T^*}(z) \leq -k$ for all $z \leq z^*$ and since k was arbitrary, it is indeed the case that $\lim_{z\to 0} w(z) = \lim_{z\to 0}\hat{w}_\infty(z) = -\infty$.

■

Proof of Proposition 2:

$$\begin{aligned}
P(D_0) &= (1+D_0)E\left[\int_0^\infty e^{-\rho s}\left(\frac{1+D_0}{1+D_t}\right)^{\gamma-1} ds\right] \\
&= (1+D_0)^\gamma \int_{-\infty}^\infty \int_0^\infty e^{-\rho s}\left(\frac{1}{1+D_0e^y}\right)^{\gamma-1} \frac{e^{-\frac{(y-\mu s)^2}{2\sigma^2 s}}}{\sqrt{2\pi\sigma^2 s}}\, ds\, dy \\
&= (1+D_0)^\gamma \int_{-\infty}^\infty \left(\frac{1}{1+D_0e^y}\right)^{\gamma-1} \frac{e^{\frac{\mu y - q|y|}{\sigma^2}}}{q}\, dy \\
&= (1+D_0)^\gamma \left[\int_{-\infty}^0 \left(\frac{1}{1+D_0e^y}\right)^{\gamma-1}\frac{e^{y\kappa}}{q}\, dy + \int_0^\infty \left(\frac{1}{1+D_0e^y}\right)^{\gamma-1}\frac{e^{y\left(\kappa-2q/\sigma^2\right)}}{q}\, dy\right] \\
&= (1+D_0)^\gamma \frac{D_0^{-\kappa}}{q}\left[V(D_0,\kappa,2-\gamma) + D_0^{\frac{2q}{\sigma^2}} V\left(\frac{1}{D_0}, \alpha + \frac{2q}{\sigma^2} - 1, 2-\gamma\right)\right].
\end{aligned}$$

In the last step we used the transformation $t = D_0e^y$ for the first integral. For the second integral, we rewrote $\left(\frac{1}{1+D_0e^y}\right)^{\gamma-1} = \left(\frac{D_0^{-1}e^{-y}}{D_0^{-1}e^{-y}+1}\right)^{\gamma-1}$ and then used the transformation $t = D_0^{-1}e^{-y}$ to get the expression. ■

Proof of Proposition 3:

We first study the case when $\alpha > 1$. We look at $\frac{P(D)}{(1+D)^{\alpha}}$, for large D. From (11), it follows that

$$\begin{aligned}\frac{P(D)}{(1+D)^{\gamma-\kappa}} &= \left(\frac{1+D}{D}\right)^{\kappa}\frac{1}{q}\left[V(D,\kappa,2-\gamma)+D^{\frac{2q}{\sigma^2}}V\left(\frac{1}{D},\alpha+\frac{2q}{\sigma^2}-1,2-\gamma\right)\right]\\ &= \frac{1+o(1)}{q}\left[\int_0^D t^{\kappa-1}(1+t)^{1-\gamma}\,dt+D^{\frac{2q}{\sigma^2}}\int_0^{1/D}t^{\alpha+\frac{2q}{\sigma^2}-2}(1+t)^{1-\gamma}\,dt\right]. \quad (24)\end{aligned}$$

Here, $\lim_{D\to\infty} o(1) = 0$. Since $\kappa > 0$ and $\gamma - \kappa > 1$, $\lim_{D\to\infty}\int_0^D t^{\kappa-1}(1+t)^{1-\gamma}\,dt = c_1$, where $0 < c_1 < \infty$. Moreover,

$$D^{\frac{2q}{\sigma^2}}\int_0^{1/D}t^{\alpha+\frac{2q}{\sigma^2}-2}(1+t)^{1-\gamma}\,dt = D^{\frac{2q}{\sigma^2}}\int_0^{1/D}t^{\alpha+\frac{2q}{\sigma^2}-2}\,dt = c_2 D^{\frac{2q}{\sigma^2}}D^{-\left(\alpha+\frac{2q}{\sigma^2}-1\right)} = c_2 D^{1-\alpha},$$

which converges to zero for large D. The finiteness of the integral is ensured, since $\alpha + \frac{2q}{\sigma^2} - 2 > -1$.
Thus, for large D, the expression converges to $\frac{c_1}{q}$.

For $\alpha < 1$, we use that

$$\begin{aligned}\frac{P(D)}{1+D} &= (1+D)^{\gamma-1}\frac{D^{-\kappa}}{q}\left[V(D,\kappa,2-\gamma)+D^{\frac{2q}{\sigma^2}}V\left(\frac{1}{D},\alpha+\frac{2q}{\sigma^2}-1,2-\gamma\right)\right]\\ &= \frac{1+o(1)}{q}D^{\gamma-1-\kappa}\left[\int_0^D t^{\kappa-1}(1+t)^{1-\gamma}\,dt+D^{\frac{2q}{\sigma^2}}\int_0^{1/D}t^{\alpha+\frac{2q}{\sigma^2}-2}(1+t)^{1-\gamma}\,dt\right].\end{aligned}$$

For the first term, we note that

$$\int_0^D t^{\kappa-1}(1+t)^{1-\gamma}\,dt = \frac{D^{\kappa}}{\kappa}\,{}_2F_1(\gamma-1,\kappa,1+\kappa,-D) = \frac{D^{\kappa-\gamma+1}}{\kappa}\,{}_2F_1\left(\gamma-1,1,1+\kappa,\frac{D}{D+1}\right).$$

For large D, the first term therefore converges to

$$\frac{1}{q\kappa}\,{}_2F_1\left(\gamma-1,1,1+\kappa,1\right) = \frac{\Gamma(\kappa+1)\Gamma(1+\kappa-\gamma)}{\Gamma(2+\kappa-\gamma)\Gamma(\kappa)} = \frac{1}{q(1-\gamma+\kappa)} = -\frac{\sigma^2}{q}\times\frac{1}{(\gamma-1)\sigma^2-\mu-q}.$$

For the second term, we note that

$$\int_0^{1/D}t^{\alpha+\frac{2q}{\sigma^2}-2}(1+t)^{1-\gamma}\,dt = \frac{D^{1-\alpha-\frac{2q}{\sigma^2}}\sigma^2}{2q+(\alpha-1)\sigma^2}\,{}_2F_1\left(\gamma-1,\alpha+\frac{2q}{\sigma^2}-1,\alpha+\frac{2q}{\sigma^2},-\frac{1}{D}\right).$$

Since ${}_2F_1\left(\gamma-1,\alpha+\frac{2q}{\sigma^2}-1,\alpha+\frac{2q}{\sigma^2},0\right) = 1$, and $\alpha = \gamma - \kappa$, the second term therefore converges to

$$\frac{\sigma^2}{q(2q+(\alpha-1)\sigma^2)} = \frac{\sigma^2}{q}\times\frac{1}{(\gamma-1)\sigma^2-\mu+q}.$$

Thus,

$$\begin{aligned}
\lim_{D\to\infty}\frac{P(D)}{1+D} &= \frac{\sigma^2}{q}\times\left(\frac{1}{(\gamma-1)\sigma^2-\mu+q}-\frac{1}{(\gamma-1)\sigma^2-\mu-q}\right)\\
&= \frac{\sigma^2}{q}\times\frac{2q}{((\gamma-1)\sigma^2-\mu)^2-q^2}\\
&= \frac{1}{\rho+\mu(\gamma-1)-(\gamma-1)^2\frac{\sigma^2}{2}}\\
&= \frac{1}{\eta}.
\end{aligned}$$

We are done.

■

Proof of Proposition 4

It is easy to see from (24) of Proposition 3 that for large D, when $\alpha > 1$, $\frac{d}{dD}\left[\frac{P(D)}{(1+D)^\alpha}\right]$ converges to 0, as does $\frac{d^2}{dD^2}\left[\frac{P(D)}{(1+D)^\alpha}\right]$. Therefore, in this case, $P' = \alpha(1+o(1))c_2D^{\alpha-1}$, and $P'' = \alpha(\alpha-1)(1+o(1))c_2D^{\alpha-2}$ for large D, and it follows that $\frac{P'(D)D}{P(D)}$ converges to α and $\frac{P''(D)D^2}{P(D)}$ converges to $\alpha(\alpha-1)$. (ii) then follows from standard Itô calculus.

For $\alpha < 1$, an identical argument for $\frac{P(D)}{D}$ proves (i). ■

Proof of Proposition 5:

(i) : The function $\frac{1}{(1+z)^\gamma}$ is analytic in the complex plane, $|z| < 1$, and can therefore be expanded in the power expansion

$$\frac{1}{(1+z)^\gamma} = \sum_{n=0}^{\infty}(-1)^n a_n z^n.$$

For $y < 0$, we use this expansion to get $1/\left(1+e^y\right)^\gamma = \sum_{n=0}^{\infty}(-1)^n a_n e^{ny}$, and for $y > 0$, we get a similar expansion $1/\left(1+e^y\right)^\gamma = e^{-\gamma y}\sum_{n=0}^{\infty}(-1)^n a_n e^{-ny}$.

Now, from Equation (16), it follows that

$$\begin{aligned}
\sqrt{2\pi\sigma^2\tau}\frac{P^\tau}{\left(1+e^d\right)^\gamma e^{-\rho\tau}} &= \int_{-\infty}^{\infty}\frac{e^{-(y-d-\mu\tau)^2/\left(2\sigma^2\tau\right)}}{\left(1+e^y\right)^\gamma}\,dy\\
&= \left(\int_{-\infty}^{-\epsilon}+\int_{\epsilon}^{\infty}+\int_{-\epsilon}^{\epsilon}\right)\frac{e^{-(y-d-\mu\tau)^2/\left(2\sigma^2\tau\right)}}{\left(1+e^y\right)^\gamma}\,dy\\
&= \int_{-\infty}^{-\epsilon}\frac{e^{-(y-d-\mu\tau)^2/\left(2\sigma^2\tau\right)}}{\left(1+e^y\right)^\gamma}\,dy+\int_{\epsilon}^{\infty}\frac{e^{-(y-d-\mu\tau)^2/\left(2\sigma^2\tau\right)}}{\left(1+e^y\right)^\gamma}\,dy+O(\epsilon)\\
&= \int_{-\infty}^{0}\frac{e^{-(y-\epsilon-d-\mu\tau)^2/\left(2\sigma^2\tau\right)}}{\left(1+e^{y-\epsilon}\right)^\gamma}\,dy+\int_{0}^{\infty}\frac{e^{-(y+\epsilon-d-\mu\tau)^2/\left(2\sigma^2\tau\right)}}{\left(1+e^{y+\epsilon}\right)^\gamma}\,dy\\
&\quad+O(\epsilon). \qquad (25)
\end{aligned}$$

For all $\epsilon > 0$ and $y < 0$, However, since,

$$\frac{e^{-(y-d-\epsilon-\mu\tau)^2/(2\sigma^2\tau)}}{(1+e^{y-\epsilon})^\gamma} = \sum_{n=0}^{\infty}(-1)^n a_n e^{-(y-\epsilon-d-\mu\tau)^2/(2\sigma^2)+n(y-\epsilon)} = \sum_{n=0}^{\infty}(-1)^n a_n e^{-\frac{\epsilon}{2}n} e^{(y-\epsilon-d-\mu\tau)^2/(2\sigma^2)\tau+n(y-\frac{\epsilon}{2})},$$

the first term is equal to

$$\int_{-\infty}^{0}\sum_{n=0}^{\infty}(-1)^n a_n e^{-\frac{\epsilon}{2}n} e^{-(y-\epsilon-d-\mu\tau)^2/(2\sigma^2\tau)+n(y-\frac{\epsilon}{2})}\,dy. \tag{26}$$

Now, define $g_{M,\epsilon}(y) = \sum_{n=0}^{M} a_n(-1)^n e^{-\epsilon n/2} e^{-(y-\epsilon-d-\mu\tau)^2/(2\sigma^2\tau)+n(y-\frac{\epsilon}{2})}$, $y < 0$, $M \in \mathbb{N}$, and $h_\epsilon(y) = e^{-(y-\epsilon-d-\mu\tau)^2/(2\sigma^2\tau)}$. Then, since $a_n \sim Cn^\gamma$ for large n, it is clear that $\sup_{n\geq 0} a_n e^{-\epsilon n/2} = C < \infty$. Therefore,

$$\begin{aligned}
|g_{M,\epsilon}(y)| &\leq C\sum_{n=0}^{M} e^{-(y-\epsilon-d-\mu\tau)^2/(2\sigma^2\tau)+n(y-\frac{\epsilon}{2})} \\
&\leq C\sum_{n=0}^{\infty} e^{-(y-\epsilon-d-\mu\tau)^2/(2\sigma^2\tau)+n(y-\frac{\epsilon}{2})} = C\frac{e^{-(y-\epsilon-d-\mu\tau)^2/(2\sigma^2\tau)}}{1-e^{-\epsilon/2}e^y} \\
&\leq C'_\epsilon e^{-(y-\epsilon-d-\mu\tau)^2/(2\sigma^2\tau)} = C'_\epsilon h_\epsilon(y).
\end{aligned}$$

Clearly, $\int_{-\infty}^{0} C'_\epsilon h_\epsilon(y)\,dy < \infty$, and therefore the dominated convergence theorem implies that $\int_{-\infty}^{0}\lim_{n\to} g_{M,\epsilon}(y)\,dy = \lim_{n\to\infty}\int_{-\infty}^{0} g_{M,\epsilon}(y)\,dy$, i.e.,

$$\begin{aligned}
\int_{-\infty}^{0}\frac{e^{-(y-\epsilon-d-\mu\tau)^2/(2\sigma^2\tau)}}{(1+e^{y-\epsilon})^\gamma}\,dy &= \sum_{n=0}^{\infty}\int_{-\infty}^{0}(-1)^n a_n e^{-\frac{\epsilon}{2}n} e^{-(y-\epsilon-d-\mu\tau)^2/(2\sigma^2\tau)+n(y-\frac{\epsilon}{2})}\,dy \\
&= \sum_{n=0}^{\infty}(-1)^n a_n e^{-\frac{\epsilon}{2}n}\int_{-\infty}^{0} e^{-(y-\epsilon-d-\mu\tau)^2/(2\sigma^2\tau)+n(y-\frac{\epsilon}{2})}\,dy
\end{aligned}$$

Define $F(x) = e^{x^2}\mathrm{Erfc}(x)$, where Erfc is the error function $\mathrm{Erfc}(x) = (\sqrt{\pi})^{-1}\int_z^\infty e^{-t^2}\,dt$ (see Abramowitz and Stegun (1964)). Then, since

$$\begin{aligned}
\frac{1}{\sqrt{2\pi\sigma^2\tau}}\int_{-\infty}^{0} e^{-(y-\epsilon d-\mu\tau)^2/(2\sigma^2\tau)+n(y-\epsilon/2)}\,dy &= \frac{1}{2}e^{n(\epsilon/2+d+\mu\tau)+n^2\tau\sigma^2/2}\mathrm{Erfc}\left(\frac{\epsilon+d+\mu\tau+n\tau\sigma^2}{\sqrt{2\sigma^2\tau}}\right) \\
&= \frac{e^{-n\frac{\epsilon}{2}}e^{-(\epsilon+d+\mu\tau)^2/(2\sigma^2\tau)}}{2}F\left(\frac{\epsilon+d+\mu\tau+n\tau\sigma^2}{\sqrt{2\sigma^2\tau}}\right),
\end{aligned}$$

it follows that

$$\begin{aligned}
\frac{1}{\sqrt{2\pi\sigma^2\tau}}\int_{-\infty}^{0}\frac{e^{-(y-\epsilon-d-\mu\tau)^2/(2\sigma^2\tau)}}{(1+e^{y-\epsilon})^\gamma}\,dy &= \frac{1}{2}\sum_{n=0}^{\infty}(-1)^n a_n e^{-\epsilon n} e^{-(\epsilon+d+\mu\tau)^2/(2\sigma^2\tau)}F\left(\frac{\epsilon+d+\mu\tau+n\tau\sigma^2}{\sqrt{2\sigma^2\tau}}\right) \\
&= (1+O(\epsilon))\frac{1}{2}e^{-(d+\mu\tau)^2/(2\sigma^2\tau)}\sum_{n=0}^{\infty}(-1)^n a_n e^{-\epsilon n}F\left(\frac{\epsilon+d+\mu\tau+n\tau\sigma^2}{\sqrt{2\sigma^2\tau}}\right).
\end{aligned}$$

An identical argument for the $\int_0^\infty \frac{e^{-(y+\epsilon-d-\mu\tau)^2/(2\sigma^2\tau)}}{(1+e^{y+\epsilon})^\gamma}\,dy$ term leads to

$$\begin{aligned}\frac{1}{\sqrt{2\pi\sigma^2\tau}}\int_0^\infty \frac{e^{-(y+\epsilon-d-\mu\tau)^2/(2\sigma^2\tau)}}{(1+e^{y+\epsilon})^\gamma}\,dy &= \frac{1}{2}\sum_{n=0}^\infty (-1)^n a_n e^{-\epsilon n} e^{-(\epsilon+d+\mu\tau)^2/(2\sigma^2\tau)} F\left(\frac{\epsilon-d-\mu\tau+(n+\gamma)\tau\sigma^2}{\sqrt{2\sigma^2\tau}}\right)\\ &= (1+O(\epsilon))e^{-(d+\mu\tau)^2/(2\sigma^2\tau)}\sum_{n=0}^\infty (-1)^n a_n e^{-\epsilon n} F\left(\frac{\epsilon-d-\mu\tau+(n+\gamma)\tau\sigma^2}{\sqrt{2\sigma^2\tau}}\right)\end{aligned}$$

Putting it all together in Equation (25), we get

$$\begin{aligned}P^\tau &= \frac{(1+e^d)^\gamma e^{-\rho\tau}}{\sqrt{2\pi\sigma^2\tau}}\left(\int_{-\infty}^0 \frac{e^{-(y-\epsilon-d-\mu\tau)^2/(2\sigma^2\tau)}}{(1+e^{y-\epsilon})^\gamma}\,dy + \int_0^\infty \frac{e^{-(y+\epsilon-d-\mu\tau)^2/(2\sigma^2\tau)}}{(1+e^{y+\epsilon})^\gamma}\,dy + O(\epsilon)\right)\\ &= O(\epsilon) + \frac{(1+e^d)^\gamma e^{-\rho\tau-(d+\mu\tau)^2/(2\sigma^2\tau)}}{2}\\ &\times \sum_{n=0}^\infty (-1)^n e^{-\epsilon n} a_n\left(F\left(\frac{\epsilon+d+\mu\tau+n\tau\sigma^2}{\sqrt{2\sigma^2\tau}}\right) + F\left(\frac{\epsilon-d-\mu\tau+(n+\gamma)\tau\sigma^2}{\sqrt{2\sigma^2\tau}}\right)\right),\end{aligned}$$

and thus, as $\epsilon \searrow 0$, we get convergence to Equation (19).

The formula is straightforward to use, since $F(x) \sim 1/x$ for large x. An error analysis implies that if n terms is used in the expansion, $\epsilon \sim \log(n)/n$ should be chosen.

(ii): When $\gamma = 1$, $a_n = 1$ for all n, and we can choose $\epsilon = 0$ and still apply the dominated convergence theorem in Equation (26) to get

$$\begin{aligned}P^\tau &= \frac{(1+e^d)^\gamma e^{-\rho\tau-(d+\mu\tau)^2/(2\sigma^2\tau)}}{2}\times\\ &\sum_{n=0}^\infty (-1)^n a_n\left(F\left(\frac{d+\mu\tau+n\tau\sigma^2}{\sqrt{2\sigma^2\tau}}\right) + F\left(\frac{-d-\mu\tau+(n+\gamma)\tau\sigma^2}{\sqrt{2\sigma^2\tau}}\right)\right), \qquad (27)\end{aligned}$$

(iii): For

$$\frac{d+\mu\tau}{\sigma^2\tau} = m \in \mathbb{N},$$

Equation (27) reduces to a case for which closed-form expressions exist, so

$$P^\tau = \frac{(1+e^d)e^{-\rho\tau-m^2\sigma^2\tau/2}}{2}\left(1+2\sum_{n=1}^{m-1}(-1)^n e^{n^2\sigma^2\tau/2}\right).$$

Finally, we note that Since $P^\tau = e^{-r(\tau)\tau}$, where $r(\tau)$ is the time-τ spot rate, we have

$$r(\tau) = \rho + \frac{\mu^2}{2\sigma^2} + \frac{1}{\tau}\Big(\log\left(-\frac{(1+e^d)^\gamma}{2}\right) + \frac{d^2}{2\sigma^2\tau} + \frac{d\mu}{\sigma^2} + \log(z)\Big),$$

where $z = \lim_{\epsilon\searrow 0}\sum_{n=0}^\infty (-1)^n e^{-\epsilon n} a_n\left(F\left(\frac{\epsilon+d+\mu\tau+n\tau\sigma^2}{\sqrt{2\sigma^2\tau}}\right) + F\left(\frac{\epsilon-d-\mu\tau+(n+\gamma)\tau\sigma^2}{\sqrt{2\sigma^2\tau}}\right)\right)$. ■

Proof of Proposition 6:
The result for r_s is standard. Using Feynman-Kac, we know that

$$P_t^\tau + \frac{1}{2}\sigma^2 z^2(1-z)^2 P_{zz}^\tau + \left[-\widehat{\mu}z(1-z) + 2\sigma^2 z(1-z)^2\right] P_z^\tau - \left[\rho + \gamma\widehat{\mu}(1-z) - \frac{1}{2}\gamma(\gamma+1)\sigma^2(1-z)^2\right] P^\tau = 0,$$

and since $P^\tau(\tau, z) = 1$, it is clear that $P(0,z) = 1 - \left[\rho + \gamma\widehat{\mu}(1-z) - \frac{1}{2}\gamma(\gamma+1)\sigma^2(1-z)^2\right]\tau + o(\tau)$, for small τ. Since $-\log(1-s) = s + O(s^2)$ for small s, it is clear that $r_s = \lim_{\tau\searrow 0} -\frac{\log(P^\tau)}{\tau} = \rho + \gamma\hat{\mu}(1-z) - \frac{1}{2}\gamma(\gamma+1)\sigma^2(1-z)^2$.

For r_l, we proceed as follows: We have

$$P^\tau = (1+e^d)^\gamma e^{-\rho\tau}\frac{1}{\sqrt{2\pi\sigma^2\tau}}\int_{-\infty}^{\infty}\frac{e^{-(y-\mu\tau)^2/(2\sigma^2\tau)}}{(1+e^{d+y})^\gamma}\,dy = (1+e^d)^\gamma e^{-\rho\tau}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{e^{-x^2/2}}{(1+e^d e^{x\sigma\sqrt{\tau}+\mu\tau})^\gamma}\,dx.$$

We study the behavior of $\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{e^{-x^2/2}}{(1+e^d e^{x\sigma\sqrt{\tau}+\mu\tau})^\gamma}\,dx$ for large τ. We decompose:

$$\begin{aligned}&\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{e^{-x^2/2}}{(1+e^d e^{x\sigma\sqrt{\tau}+\mu\tau})^\gamma}\,dx\\ &\quad= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-\frac{\mu\tau+d}{\sigma\sqrt{\tau}}}\frac{e^{-x^2/2}}{(1+e^{x\sigma\sqrt{\tau}+\mu\tau+d})^\gamma}\,dx + \frac{1}{\sqrt{2\pi}}\int_{-\frac{\mu\tau+d}{\sigma\sqrt{\tau}}}^{\infty}\frac{e^{-x^2/2}}{(1+e^{x\sigma\sqrt{\tau}+\mu\tau+d})^\gamma}\,dx. \qquad (28)\end{aligned}$$

We prove the results for r^l by studying the first and second term in Equation (28) separately for the two cases $\mu \le \gamma\sigma^2$ and $\mu > \gamma\sigma^2$ respectively. By showing that the first term behaves like $e^{-\frac{\mu}{2\sigma^2}\tau}$ for large τ for all μ, whereas the second term behaves like $e^{-\frac{\mu}{2\sigma^2}\tau}$ when $\mu \le \gamma\sigma^2$ and like $e^{-(\gamma\mu-\gamma^2\sigma^2/2)\tau}$ when $\mu > \gamma\sigma^2$, the result will follow.

Since $0 < e^{x\sigma\sqrt{\tau}+\mu\tau+d} \le 1$ for $x \le -\frac{\mu\tau+d}{\sigma\sqrt{\tau}}$, we have

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-\frac{\mu\tau+d}{\sigma\sqrt{\tau}}}\frac{e^{-x^2/2}}{(1+e^{x\sigma\sqrt{\tau}+\mu\tau+d})^\gamma}\,dx = C\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-\frac{\mu\tau+d}{\sigma\sqrt{\tau}}} e^{-x^2/2}\,dx = C\times N\left(-\frac{\mu\tau+d}{\sigma\sqrt{\tau}}\right),$$

for some $C \in [1/2^\gamma, 1]$, where $N(\cdot)$ is the cumulative normal distribution function, $N(v) \stackrel{\text{def}}{=} \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{v} e^{-y^2/2}\,dy$. Now, we use

$$N(-v) = C_2\frac{e^{-v^2/2}}{v}, \qquad C_2 \in \frac{1}{\sqrt{2\pi}}\left[\frac{v^2}{1+v^2}, 1\right], \qquad (29)$$

which is valid for $v \gg 0$, to get

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-\frac{\mu\tau+d}{\sigma\sqrt{\tau}}}\frac{e^{-x^2/2}}{(1+e^{x\sigma\sqrt{\tau}+\mu\tau+d})^\gamma}\,dx = C\times C_2\frac{e^{-q^2/2}}{q} = C_3\frac{e^{-\frac{\mu^2}{2\sigma^2}\tau-\frac{\mu d}{\sigma^2}-\frac{d^2}{2\sigma^2\tau}}}{q},$$

where

$$C_3 \in \frac{1}{\sqrt{2\pi}}\left[\frac{1}{2^{\gamma+1}}, 1\right], \qquad \text{and} \qquad q = \frac{\mu\tau+d}{\sigma\sqrt{\tau}}.$$

We next study the second term in Equation (28), when $\mu < \gamma\sigma^2$. First, we note that $\mu < \gamma\sigma^2$ implies that $\gamma\sigma - \frac{\mu}{\sigma} > 0$. Obviously, $\frac{1}{(1+e^{x\sigma\sqrt{\tau}+\mu\tau+d})^\gamma} \le e^{-\gamma(x\sigma\sqrt{\tau}+\mu\tau+d)}$, so

$$
\begin{aligned}
0 \le \frac{1}{\sqrt{2\pi}}\int_{-\frac{\mu\tau+d}{\sigma\sqrt{\tau}}}^{\infty} \frac{e^{-x^2/2}}{(1+e^{x\sigma\sqrt{\tau}+\mu\tau+d})^\gamma}\,dx &\le \frac{1}{\sqrt{2\pi}}\int_{-\frac{\mu\tau+d}{\sigma\sqrt{\tau}}}^{\infty} e^{-(x^2+2x\gamma\sigma\sqrt{\tau})/2-\gamma\mu\tau-\gamma d}\,dx \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\frac{\mu\tau+d}{\sigma\sqrt{\tau}}}^{\infty} e^{-(x+\gamma\sigma\sqrt{\tau})^2/2+\frac{\gamma^2\sigma^2\tau}{2}-\gamma\mu\tau-d\gamma}\,dx \\
&= e^{-\gamma d}e^{-\tau(\gamma\mu-\gamma^2\sigma^2/2)}\frac{1}{\sqrt{2\pi}}\int_{-\frac{\mu\tau+d}{\sigma\sqrt{\tau}}+\gamma\sigma\sqrt{\tau}}^{\infty} e^{-x^2/2}\,dx \\
&= e^{-\gamma d}e^{-\tau(\gamma\mu-\gamma^2\sigma^2/2)}\frac{1}{\sqrt{2\pi}}\int_{\left(\gamma\sigma-\frac{\mu}{\sigma}\right)\sqrt{\tau}-\frac{d}{\sigma\sqrt{\tau}}}^{\infty} e^{-x^2/2}\,dx \\
&= e^{-\gamma d}e^{-\tau(\gamma\mu-\gamma^2\sigma^2/2)}N\left(-\left(\gamma\sigma-\frac{\mu}{\sigma}\right)\sqrt{\tau}+\frac{d}{\sigma\sqrt{\tau}}\right) \\
&\le e^{-\gamma d}e^{-\tau(\gamma\mu-\gamma^2\sigma^2/2)}\frac{1}{\sqrt{2\pi}}\frac{e^{-q_2^2/2}}{q_2} \\
&= e^{-\gamma d}e^{-\tau(\gamma\mu-\gamma^2\sigma^2/2)}\frac{1}{\sqrt{2\pi}}\frac{e^{-\frac{d^2}{2\sigma^2\tau}+\gamma d-\frac{d\mu}{\sigma^2}-(\gamma^2\frac{\sigma^2}{2}-\gamma\mu+\frac{\mu^2}{2\sigma^2})\tau}}{q_2} \\
&= e^{-\frac{d^2}{2\sigma^2\tau}-\frac{d\mu}{\sigma^2}} \times \frac{1}{\sqrt{2\pi}}\frac{e^{\frac{-\mu^2}{2\sigma^2}\tau}}{q_2},
\end{aligned}
$$

where $q_2 = \left(\gamma\sigma - \frac{\mu}{\sigma}\right)\sqrt{\tau} - \frac{d}{\sigma\sqrt{\tau}}$, and we used that $\frac{1}{\sqrt{2\pi}}\int_v^\infty e^{-y^2/2}\,dy = N(-v)$, and Equation (29). Thus,

$$
\frac{1}{\sqrt{2\pi}}\int_{-\frac{\mu\tau+d}{\sigma\sqrt{\tau}}}^{\infty} \frac{e^{-x^2/2}}{(1+e^{x\sigma\sqrt{\tau}+\mu\tau+d})^\gamma}\,dx = C_4 e^{-\frac{d^2}{2\sigma^2\tau}-\frac{d\mu}{\sigma^2}} \times \frac{e^{\frac{-\mu^2}{2\sigma^2}\tau}}{q_2},
$$

where $C_4 \in \left[0, \frac{1}{\sqrt{2\pi}}\right]$. Putting it all together, for large τ we get

$$
\begin{aligned}
P^\tau &= (1+e^d)^\gamma e^{-\rho t}\left(\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-\frac{\mu\tau+d}{\sigma\sqrt{\tau}}} \frac{e^{-x^2/2}}{(1+e^{x\sigma\sqrt{\tau}+\mu\tau+d})^\gamma}\,dx + \frac{1}{\sqrt{2\pi}}\int_{-\frac{\mu\tau+d}{\sigma\sqrt{\tau}}}^{\infty} \frac{e^{-x^2/2}}{(1+e^{x\sigma\sqrt{\tau}+\mu\tau+d})^\gamma}\,dx\right) \\
&= (1+e^d)^\gamma e^{-\rho t}\left(C_3\frac{e^{-\frac{\mu^2}{2\sigma^2}\tau-\frac{\mu d}{\sigma^2}-\frac{d^2}{2\sigma^2\tau}}}{q} + C_4 e^{-\frac{d^2}{2\sigma^2\tau}-\frac{d\mu}{\sigma^2}} \times \frac{e^{\frac{-\mu^2}{2\sigma^2}\tau}}{q_2}\right) \\
&= e^{-\left(\rho+\frac{\mu^2}{2\sigma^2}\right)\tau}(1+e^d)^\gamma e^{-\frac{\mu d}{\sigma^2}-\frac{d^2}{2\sigma^2\tau}}\left(\frac{C_3}{q}+\frac{C_4}{q_2}\right).
\end{aligned}
$$

Therefore,

$$
-\frac{\log(P^\tau)}{\tau} = \rho + \frac{\mu^2}{2\sigma^2} + \frac{Q(\tau)}{\tau}, \qquad \text{where } Q(\tau) = \log\left((1+e^d)^\gamma e^{-\frac{\mu d}{\sigma^2}-\frac{d^2}{2\sigma^2\tau}}\left(\frac{C_3}{q}+\frac{C_4}{q_2}\right)\right).
$$

Now, $Q(\tau) = \log\left((1+e^d)^\gamma\right) - \frac{\mu d}{\sigma^2} - \frac{d^2}{2\sigma^2\tau} + \log\left(\frac{C_3}{q}+\frac{C_4}{q_2}\right)$, and since $C_3 \in \frac{1}{\sqrt{2\pi}}\left[\frac{1}{2^{\gamma+1}}, 1\right]$, $C_4 \in$

$\left[0, \frac{1}{\sqrt{2\pi}}\right]$, $q = \frac{\mu\tau+d}{\sigma\sqrt{\tau}}$ and $q_2 = \left(\gamma\sigma - \frac{\mu}{\sigma}\right)\sqrt{\tau} - \frac{d}{\sigma\sqrt{\tau}}$, it follows that $|Q(\tau)| = o(\tau)$ for large τ, i.e., that $\lim_{\tau\to\infty} \frac{|Q(\tau)|}{\tau} = 0$. From this it immediately follows that $\lim_{\tau\to\infty} -\frac{\log(P^\tau)}{\tau} = \rho + \frac{\mu^2}{2\sigma^2}$.

We now consider the case when $\mu > \gamma\sigma^2$, and define $v = \mu/\sigma - \gamma\sigma > 0$. We first note that $\frac{\mu^2}{2\sigma^2} \geq \gamma\mu - \gamma^2\sigma^2/2$, since $\mu^2/(2\sigma^2) - \gamma\mu + \gamma^2\sigma^2/2 = \frac{1}{2\sigma^2}(\mu - \gamma\sigma^2)^2 \geq 0$. Thus, since the $\int_{-\infty}^{-\frac{\mu\tau+d}{\sigma\sqrt{\tau}}} \frac{e^{-x^2/2}}{(1+e^d e^{x\sigma\sqrt{\tau}+\mu\tau})^\gamma}\, dx$-term in Equation (28) behaves like $e^{-\tau\times\mu^2/(2\sigma^2)}$ for large τ, if the $\int_{-\frac{\mu\tau+d}{\sigma\sqrt{\tau}}}^{\infty} \frac{e^{-x^2/2}}{(1+e^{x\sigma\sqrt{\tau}+\mu\tau+d})^\gamma}\, dx \sim e^{-\tau(\mu\gamma-\gamma^2\sigma^2/2)}$, for large τ, then the result we wish to prove follows, since it is always the case that $c_1 e^{-\alpha_1\tau} + c_2 e^{-\alpha_2\tau} \sim e^{-\min(\alpha_1,\alpha_2)\tau}$ for large τ, for arbitrary $c_1 > 0$, $c_2 > 0$, $\alpha_1 > 0$, $\alpha_2 > 0$.

We have

$$\begin{aligned}
\frac{1}{\sqrt{2\pi}} \int_{-\frac{\mu\tau+d}{\sigma\sqrt{\tau}}}^{\infty} \frac{e^{-x^2/2}}{(1+e^{x\sigma\sqrt{\tau}+\mu\tau+d})^\gamma}\, dx &\leq \frac{1}{\sqrt{2\pi}} \int_{-\frac{\mu\tau+d}{\sigma\sqrt{\tau}}}^{\infty} e^{-(x^2+2x\gamma\sigma\sqrt{\tau})/2-\gamma\mu\tau-\gamma d}\, dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\frac{\mu\tau+d}{\sigma\sqrt{\tau}}}^{\infty} e^{-(x+\gamma\sigma\sqrt{\tau})^2/2+\frac{\gamma^2\sigma^2\tau}{2}-\gamma\mu\tau-d\gamma}\, dx \\
&= e^{-\gamma d} e^{-\tau(\gamma\mu-\gamma^2\sigma^2/2)} \frac{1}{\sqrt{2\pi}} \int_{-\frac{\mu\tau+d}{\sigma\sqrt{\tau}}+\gamma\sigma\sqrt{\tau}}^{\infty} e^{-x^2/2}\, dx \\
&= e^{-\gamma d} e^{-\tau(\gamma\mu-\gamma^2\sigma^2/2)} N\left(v\sqrt{\tau} + \frac{d}{\sigma\sqrt{\tau}}\right) \\
&= e^{-\gamma d} e^{-\tau(\gamma\mu-\gamma^2\sigma^2/2)}(1 - O(e^{-v\tau})).
\end{aligned}$$

Also, since $1 + e^{x\sigma\sqrt{\tau}+\mu\tau+d} \leq 2e^{x\sigma\sqrt{\tau}+\mu\tau+d}$

$$\begin{aligned}
\frac{1}{\sqrt{2\pi}} \int_{-\frac{\mu\tau+d}{\sigma\sqrt{\tau}}}^{\infty} \frac{e^{-x^2/2}}{(1+e^{x\sigma\sqrt{\tau}+\mu\tau+d})^\gamma}\, dx &\geq \frac{1}{2^\gamma} \frac{1}{\sqrt{2\pi}} \int_{-\frac{\mu\tau+d}{\sigma\sqrt{\tau}}}^{\infty} e^{-(x^2+2x\gamma\sigma\sqrt{\tau})/2-\gamma\mu\tau-\gamma d}\, dx \\
&= \frac{1}{2^\gamma} \frac{1}{\sqrt{2\pi}} \int_{-\frac{\mu\tau+d}{\sigma\sqrt{\tau}}}^{\infty} e^{-(x+\gamma\sigma\sqrt{\tau})^2/2+\frac{\gamma^2\sigma^2\tau}{2}-\gamma\mu\tau-d\gamma}\, dx \\
&= \frac{1}{2^\gamma} e^{-\gamma d} e^{-\tau(\gamma\mu-\gamma^2\sigma^2/2)} \frac{1}{\sqrt{2\pi}} \int_{-\frac{\mu\tau+d}{\sigma\sqrt{\tau}}+\gamma\sigma\sqrt{\tau}}^{\infty} e^{-x^2/2}\, dx \\
&= \frac{1}{2^\gamma} e^{-\gamma d} e^{-\tau(\gamma\mu-\gamma^2\sigma^2/2)} N\left(v\sqrt{\tau} + \frac{d}{\sigma\sqrt{\tau}}\right) \\
&= \frac{1}{2^\gamma} e^{-\gamma d} e^{-\tau(\gamma\mu-\gamma^2\sigma^2/2)}(1 - O(e^{-v\tau})).
\end{aligned}$$

Thus, it is the case that

$$\frac{1}{\sqrt{2\pi}} \int_{-\frac{\mu\tau+d}{\sigma\sqrt{\tau}}}^{\infty} \frac{e^{-x^2/2}}{(1+e^{x\sigma\sqrt{\tau}+\mu\tau+d})^\gamma}\, dx = C_5 e^{-\tau(\gamma\mu-\gamma^2\sigma^2/2)},$$

where $C_5 \in \left[\frac{e^{-\gamma d}}{2^\gamma} - \epsilon, e^{-\gamma d} + \epsilon\right]$, for arbitrary $\epsilon > 0$, for large enough τ.

We therefore get

$$-\frac{\log(P^{\tau})}{\tau} = -\frac{1}{\tau}\log\left((1+e^{d})^{\gamma}e^{-\rho\tau}\left(e^{-\tau\frac{\mu^2}{2\sigma^2}}e^{-\frac{\mu d}{\sigma^2}-\frac{d^2}{2\sigma^2\tau}}\frac{C_3}{q}+C_5e^{-\tau(\gamma\mu-\gamma^2\sigma^2/2)}\right)\right).$$

Now, since $\frac{\mu^2}{2\sigma^2} \geq \gamma\mu - \gamma^2\sigma^2/2$, the second term within the log expression dominates the first, so we get

$$-\frac{\log(P^{\tau})}{\tau} = -\frac{1}{\tau}\left(\log\left((1+e^{d})^{\gamma}e^{-\rho\tau}C_5e^{-\tau(\gamma\mu-\gamma^2\sigma^2/2)}\right)+o(\tau)\right) = \frac{(\rho+\gamma\mu-\gamma^2\sigma^2/2)\tau+o(\tau)}{\tau},$$

so indeed $\lim_{\tau\to\infty} -\frac{\log(P^{\tau})}{\tau} = \rho+\gamma\mu-\gamma^2\sigma^2/2 = \rho+\gamma(\mu+\sigma^2/2)-\gamma(\gamma+1)\sigma^2/2$. ■

Proof of Proposition 7:

Without loss of generality, we assume that $\underline{s} \leq \overline{s}$, since the whole proof otherwise goes through by replacing $\underline{s}$ with $\overline{s}$.

We begin with (ii): It is easy to show the following inequality, which is valid for an arbitrary constant, $x \leq 0$:

$$\int_{\bar{t}}^{\infty} e^{-\rho s}\frac{e^{-\frac{(x-\mu s)^2}{2\sigma^2 s}}}{\sqrt{2\pi\sigma^2 s}}\,ds \geq \frac{e^{-(\rho+\mu\bar{t})}}{q}e^{\kappa x}, \tag{30}$$

where κ and q are defined in (4).

Now,

$$
\begin{aligned}
P(C_0) = E\left[\int_0^\infty e^{-\rho t}\left(\frac{f(D_0)}{f(D_t)}\right)^\gamma f(D_t)\,dt\right] &\geq f(D_0)^\gamma E\left[\int_{\bar{t}}^\infty e^{-\rho t} f(D_t)^{1-\gamma}\,dt\right] \\
&\geq c_0^\gamma D_0^\gamma \int_{\bar{t}}^\infty e^{-\rho t} E\left[f(D_t)^{1-\gamma}\right]\,dt \\
&\geq c_0^\gamma D_0^\gamma \int_{\bar{t}}^\infty e^{-\rho t} E\left[f(D_t)^{1-\gamma} I_{s_t \leq \underline{s}}\right]\,dt \\
&\geq c_0^\gamma D_0^\gamma f(e^{\underline{s}})^{1-\gamma} \int_{\bar{t}}^\infty e^{-\rho t} E\left[I_{s_t \leq \underline{s}}\right]\,dt \\
&\geq c_0^\gamma D_0^\gamma f(e^{\underline{s}})^{1-\gamma} \int_{\bar{t}}^\infty e^{-\rho t} \Phi\left(\frac{\underline{s} - s_0 - \mu t}{\sigma\sqrt{t}}\right)\,dt \\
&= c_0^\gamma D_0^\gamma f(e^{\underline{s}})^{1-\gamma} \int_{\bar{t}}^\infty \int_{-\infty}^{\underline{s}-s_0} e^{-\rho t} \frac{e^{-\frac{(x-\mu t)^2}{2\sigma^2 t}}}{\sqrt{2\pi\sigma^2 t}}\,dx\,dt \\
&= c_0^\gamma D_0^\gamma f(e^{\underline{s}})^{1-\gamma} \int_{-\infty}^{\underline{s}-s_0} \int_{\bar{t}}^\infty e^{-\rho t} \frac{e^{-\frac{(x-\mu t)^2}{2\sigma^2 t}}}{\sqrt{2\pi\sigma^2 t}}\,dt\,dx \\
&\geq c_0^\gamma D_0^\gamma f(e^{\underline{s}})^{1-\gamma} \frac{e^{-(\rho+\mu\bar{t})}}{q} \int_{-\infty}^{\underline{s}-s_0} e^{\kappa x}\,dx \\
&= c_0^\gamma D_0^\gamma f(e^{\underline{s}})^{1-\gamma} \frac{e^{-(\rho+\mu\bar{t})}}{q} \frac{e^{\kappa\underline{s}}}{\kappa} \times e^{-\kappa s_0} \\
&= c_0^\gamma D_0^\gamma f(e^{\underline{s}})^{1-\gamma} \frac{e^{-(\rho+\mu\bar{t})}}{q} \frac{e^{\kappa\underline{s}}}{\kappa} \times D_0^{-\kappa} \\
&= c_0^\gamma f(e^{\underline{s}})^{1-\gamma} \frac{e^{-(\rho+\mu\bar{t})}}{q} \frac{e^{\kappa\underline{s}}}{\kappa} \times D_0^{\alpha} \\
&\geq c_0^\gamma f(e^{\underline{s}})^{1-\gamma} \frac{e^{-(\rho+\mu\bar{t})}}{q} \frac{e^{\kappa\underline{s}}}{\kappa} c_1^{-\alpha} \times f(D_0)^{\alpha} \\
&= c_3 C_0^\alpha.
\end{aligned}
$$

For (i), we note that when $f(\epsilon) < c_2\epsilon$, we can choose an arbitrary $m > \max\{0, -\underline{s}\}$, to bound

$$
\begin{aligned}
P(C_0) = E\left[\int_0^\infty e^{-\rho t}\left(\frac{f(D_0)}{f(D_t)}\right)^\gamma f(D_t)\,dt\right] &\geq f(D_0)^\gamma E\left[\int_{\bar{t}}^\infty e^{-\rho t} f(D_t)^{1-\gamma}\,dt\right] \\
&\geq c_0^\gamma D_0^\gamma \int_{\bar{t}}^\infty e^{-\rho t} E\left[f(D_t)^{1-\gamma}\right]\,dt \\
&\geq c_0^\gamma D_0^\gamma \int_{\bar{t}}^\infty e^{-\rho t} E\left[f(D_t)^{1-\gamma} I_{s_t \leq -m}\right]\,dt \\
&\geq c_0^\gamma D_0^\gamma f(e^{-m})^{1-\gamma} \int_{\bar{t}}^\infty e^{-\rho t} E\left[I_{s_t \leq -m}\right]\,dt \\
&\geq c_0^\gamma D_0^\gamma f(e^{-m})^{1-\gamma} \int_{\bar{t}}^\infty e^{-\rho t} \Phi\left(\frac{-m - s_0 - \mu t}{\sigma\sqrt{t}}\right)\,dt \\
&= c_0^\gamma D_0^\gamma f(e^{-m})^{1-\gamma} \int_{\bar{t}}^\infty \int_{-\infty}^{-m-s_0} e^{-\rho t} \frac{e^{-\frac{(x-\mu t)^2}{2\sigma^2 t}}}{\sqrt{2\pi\sigma^2 t}}\,dx\,dt \\
&= c_0^\gamma D_0^\gamma f(e^{-m})^{1-\gamma} \int_{-\infty}^{-m-s_0} \int_{\bar{t}}^\infty e^{-\rho t} \frac{e^{-\frac{(x-\mu t)^2}{2\sigma^2 t}}}{\sqrt{2\pi\sigma^2 t}}\,dt\,dx \\
&\geq c_0^\gamma D_0^\gamma f(e^{-m})^{1-\gamma} \frac{e^{-(\rho+\mu\bar{t})}}{q} \int_{-\infty}^{-m-s_0} e^{\kappa x}\,dx \\
&= c_0^\gamma D_0^\gamma f(e^{-m})^{1-\gamma} \frac{e^{-(\rho+\mu\bar{t})}}{q} \frac{e^{-\kappa m}}{\kappa} \times e^{-\kappa s_0} \\
&= c_0^\gamma D_0^\gamma f(e^{-m})^{1-\gamma} \frac{e^{-(\rho+\mu\bar{t})}}{q} \frac{e^{-\kappa m}}{\kappa} \times D_0^{-\kappa} \\
&= c_0^\gamma f(e^{-m})^{1-\gamma} \frac{e^{-(\rho+\mu\bar{t})}}{q} \frac{e^{-\kappa m}}{\kappa} \\
&\geq c_0^\gamma c_2^{1-\gamma} \frac{e^{-(\rho+\mu\bar{t})}}{q} \frac{1}{\kappa} D_0^\alpha \times e^{m(\gamma-\kappa-1)} \\
&= c_4(D_0) e^{m(\alpha-1)}.
\end{aligned}
$$

Now, since $\alpha > 1$ and m is arbitrary, there $P(C_0)$ must be infinite, and the equilibrium does therefore not exist. Equivalently, we could have used the identity $\frac{1}{1-\gamma}\frac{P(C_0)}{C_0} = U$ to show that expected utility is negative infinity for this case. ■

Mathematica code

Price-dividend ratios

We have verified numerically that the formulae for the prices given in Proposition 2 are indeed correct, both above and below the breakpoint. The following Mathematica code calculates the price-dividend ratios for different D, for a long, but finite horizon, economy ($T = 1000$), using direct numerical integration of (10), and produces identical results as the ones shown in Figure 1.

```
In[1]:= γ = 5; σ = 4/100; μ = 0.75/100; ρ = 1/100; ξ = μ − σ^2/2; T = 1000; B = 1; PD={};
In[2]:= v=Range[1/4,8,1/4];
In[3]:= For[i=1,i<32,
     e=Extract[v,i];
     v=NIntegrate[((B + e)/(B + e ∗ Exp[y]))^(γ−1) ∗ Exp[−ρ ∗ τ − (y − ξ ∗ τ)^2/(2 ∗ σ^2 ∗ τ)]
    /Sqrt[2 ∗ π ∗ σ^2 ∗ τ], {y, −∞, ∞}, {τ, 0, T}];
     PD=Append[PD,{e,v}],
     i=i+1];
In[4]:= ListPlot[PD,PlotJoined->True,PlotRange->All];
```

Long-term risk-free rate

We have verified numerically that the formulae for the long rate given in Proposition 6 are indeed correct, by directly evaluating Equation (16). The following Mathematica code calculates the yield for different maturities.

For example, with parameters, $\rho = 1\%$, $\mu = 3.5\%$, $\sigma = 20\%$, $\gamma = 2.5$, the long rate is close to $r^l = \rho + \frac{\mu^2}{2\sigma^2} = 2.53\%$ in line with Equation (20). The list L provides pairs of time to maturity and yields, $\{t, r_t\}$. For example, the last element in L shows that for a time to maturity of 10,000 years the yield is 2.56% in this example.

By varying $B0$, $D0$ and γ in the code, it is easily verified that the long rate does not depend on these parameters. It can also be checked that, for $\mu > \gamma\sigma^2$, Equation (21) provides the correct long rate.

```
In[1]:= B0 = 2; D0 = 1; σ = 0.2; μ = 0.035; γ = 2.5; ρ = 0.01;Off[Integrate::gener];
In[2]:= L = {}; T = {1, 10, 100, 1000, 10000, −1};
In[3]:= For[ t = First[T], t > 0,
    P= N[Integrate[(B0 + D0)^γ ∗ Exp[−ρt]∗
     1/Sqrt[2π ∗ σ^2 t] ∗ Exp[−(y − μt)^2/(2 ∗ σ^2 t)]/(B0 + D0 ∗ Exp[y])^γ, {y, −∞, ∞}]];
    r = −Log[P]/t;
    L = Append[L, {t, r}]; T = Delete[T, 1]; t= First[T];]

In[4]:= L (* L is a list with elements {t, r_t}, from numerical calculations*)
Out[4]= {{1,0.0362381},{10,0.0350963},{100,0.0307781},{1000,0.026798},{10000,0.0255731}}

In[5]:= r_l = If[μ < γσ^2, ρ + μ^2/(2σ^2), ρ + γμ − γ^2σ^2/2] (* Theoretical value of long rate *)
Out[5]= 0.0253125
```

References

Abel, A. (1990): "Asset Prices under Habit Formation and Catching up with the Joneses," *American Economic Review*, pp. 38–42.

Abramowitz, M., and I. A. Stegun (1964): *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* Dover, New York.

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

Barro, R. (2005): "Rare Events and the Equity Premium," Working paper, Harvard University, Department of Economics.

Berk, J., and J. Walden (2009): "Limited Capital Market Participation and Human Capital Risk," Working paper, Stanford GSB.

Blanchard, O. J. (1979): "Speculative Bubbles, Crashes and Rational Expectations," *Economics Letters*, 3, 387–389.

Blanchard, O. J., and M. W. Watson (1982): "Bubbles, Rational Expectations and Financial Markets," in *Crises in the Economic and Financial Structure*, ed. by P. Wachtel, pp. 295–315. Lexington Books, Lexington.

Campbell, J. Y. (1986): "Bond and Stock Returns in a Simple Exchange Model," *Quarterly Journal of Economics*, pp. 785–803.

Campbell, J. Y. (2003): "Consumption-Based Asset Pricing," in *Handbook of the Economics of Finance*, ed. by G. M. Constantinides, M. Harris, and R. M. Stulz, pp. 805–887. Elsevier.

Campbell, J. Y., and J. H. Cochrane (1999): "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107, 205–251.

Cochrane, J. H. (1992): "Explaining the Variance of Price-Dividend Ratios," *Review of Financial Studies*, 5, 243–280.

Cochrane, J. H., F. A. Longstaff, and P. Santa-Clara (2008): "Two Trees," *Review of Financial Studies*, 21(1), 347–385.

Diba, B. T., and H. I. Grossman (1988): "Explosive Rational Bubbles in Stock Prices?," *American Economic Review*, 78, 520–530.

Dybvig, P. H., J. E. Ingersoll, Jr., and S. A. Ross (1996): "Long Forward and Zero-Coupon Rates Can Never Fall," *Journal of Business*, 69, 1–25.

Engsted, T. (2006): "Explosive Bubbles in the Cointegrated VAR Model," *Finance Research Letters*, 3, 154–162.

Froot, K. A., and M. Obstfeld (1991): "Intrinsic Bubbles: The Case of Stock Prices," *American Economic Review*, 81, 1189–1214.

Gabaix, X., and D. Laibson (2001): "The 6D Bias and the Equity Premium Puzzle," in *NBER Macroeconomics Annual*, ed. by B. Bernanke, and K. Rogoff. MIT Press.

Geweke, J. (2001): "A Note on Some Limitations of CRRA Utility," *Economic Letters*, 71, 341–345.

Gradshteyn, L. S., and I. M. Ryzhik (2000): *Tables of Integrals, Series and Products*. Academic Press.

Hansen, L. P., and R. Jagannathan (1991): "Implications of Security Market Data for Models of Dynamic Economies," *Journal of Political Economy*, 99(2), 225–262.

Ingersoll, Jr., J. E. (1987): *Theory of Financial Decision Making*. Rowman and Littlefield, Totowa, NJ.

Kim, T. S., and E. Omberg (2002): "Dynamic Nonmyopic Portfolio Optimization," *Review of Financial Studies*, 9, 141–161.

Kogan, L., S. Ross, J. Wang, and M. Westerfield (2006): "The Price Impact and Survival of Irrational Traders," *Journal of Finance*, 61(1), 195–229.

LeRoy, S., and R. Porter (1981): "The Present Value Relation: Tests Based on Variance Bounds," *Econometrica*, 49, 555–577.

LeRoy, S. F. (2006): "Excess Volatility," in *The New Palgrave Dictionary of Economics*, ed. by S. N. Durlauf, and L. E. Blume. Palgrave Macmillan, 2 edn.

Lucas, R. (1978): "Asset Prices in an Exchange Economy," *Econometrica*, 46, 1429–1445.

Martin, I. (2008): "The Lucas Orchard," Working paper (November), Stanford GSB.

Mehra, R., and E. C. Prescott (1985): "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 15, 145–161.

Parker, J. (2001): "The Consumption Risk of the Stock Market," *Brookings Papers on Economic Activity*, 2, 279–348.

Rietz, T. A. (1988): "The Equity Risk Premium: A Solution," *Journal of Monetary Economics*, 22(1), 117–131.

Rubinstein, M. (1976): "The Valuation of Uncertain Income Streams and the Pricing of Options," *Bell Journal of Economics*, 7(2), 407–425.

Shiller, R. J. (1981): "Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?," *American Economic Review*, 71, 421–436.

Wachter, J. (2002): "Portfolio and Consumption Decisions under Mean-Reverting Returns: An Exact Solution for Complete Markets," *Journal of Financial and Quantitative Analysis*, pp. 63–92.

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

Weitzman, M. L. (1998): "Why the Far-Distant Future Should be Discounted at its Lowest Possible Rate," *Journal of Environmental Economics and Management*, 36, 201–208.

——— (2001): "Gamma Discounting," *American Economic Review*, 91, 260–271.

——— (2007): "Subjective Expectations and Asset-Return Puzzles," *American Economic Review*, 97, 1102–1130.

West, K. D. (1987): "A Specification Test for Speculative Bubbles," *Quarterly Journal of Economics*, 102, 553–580.