The Long-Term Discount Rate

June, 2009

Abstract

We develop an expression for the long-term discount rate in an economy in which a representative consumer has access to both a risk-free and a risky production technology. Even when the risk-free sector is very small, and with probability one becomes a negligible fraction of the economy in the long run, interest rates are determined differently than in a single-sector economy. As in a single-risky-sector economy, the short rate depends on risk aversion; however, the long rate depends on consumption growth and volatility (i.e., the production possibilities of the economy), but *not* on the representative investor's risk aversion. Rather than always being flat, the yield curve is usually upward sloping but may be downward sloping. In the latter case, the slope predicts a future increase in consumption growth rates and lower future short-term interest rates.

1 Introduction

How should society value a \$1 certain cash flow in the distant future? Any rational investment in projects whose benefits accrue over long time horizons requires an answer to this question. Further, since traded bonds typically have maturities less than thirty years, the answer must necessarily be theoretical.

Much of our intuition concerning the behavior of the long-term discount rate comes from the one-tree model of Lucas (1978), in which long-term and short-term rates are constant, equal, and depend heavily on the level of risk aversion in the economy. In this paper, we show that this intuition is actually extremely fragile by considering a simple two-tree extension of this economy. Despite the fact that, with probability one, our model looks indistinguishable from the one-tree Lucas (1978) model in the long run, long-term interest rates behave completely differently in the two economies.[1]

We consider how a risk-averse agent who consumes dividends from two sources, one risky and one risk-free, values cash flows in the future. Dividends from the risky asset follow a geometric Brownian motion, and all dividends are consumed instantaneously. Mechanically then, our paper essentially characterizes the term structure in a Lucas economy with *two* trees, one with a risky, and the other a riskless, dividend.[2] This is the simplest possible extension of the one-tree Lucas (1978) model and, as long as the growth rate of the risky tree exceeds that of the riskless tree, this economy in the long run looks indistinguishable from a one-tree economy. Nevertheless, it has very different implications for interest rate behavior. In particular, we find that the long and short rates are determined by different variables. The short rate essentially follows the same dynamics as it would in a one-tree economy, being the sum of the personal discount rate, the expected growth rate times the risk aversion coefficient, and the precautionary savings motive; however, the long rate depends only on the growth rate and volatility of consumption. Thus, the appropriate rate at which to discount long-term cash flows only depends on the technological characteristics of the economy and does not depend on the risk-aversion of the representative agent. This result holds regardless of the fraction of the economy — strictly between zero and one — that is made up of the riskless sector. The reason for the difference between the one-tree and the two-tree models is the insurance provided by the riskless tree in extreme low probability, low consumption, states of the world, i.e., in the far-left tail of the distribution. Even if the

[1]The degeneracy of our model is important. If we started with a model that always looked different from the one-tree model, we would not then be particularly surprised to find that many economic variables, including the long rate, differed between the two economies.

[2]Building on the one-tree framework of Lucas (1978), the two-tree model was first introduced in Cochrane, Longstaff, and Santa-Clara (2008) and then further extended by Martin (2008). Further, Santos and Veronesi (2006) also present a multiple sector asset pricing economy which in equilibrium has the characteristics of the two-tree economy.

riskless tree is arbitrarily small, it significantly influences the long-term rate.

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 the two-tree 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. Finally, we observe that bond prices will be driven by the maximum consumption available to the agent. Therefore, the yields derived from the bond prices will be equivalent and will depend on the economic variables that determine the agent's consumption, not on risk aversion. This observation has implications for evaluating long term investments across countries; by the logic of this approach, different discount rates should be used in countries with different expected growth rates.

Weitzman (1998, 2001) (see also Weitzman (2005)) has pioneered investigations into the appropriate discount rate to use when valuing long-term projects such as those required to mitigate climate change. He argues that, if there is parameter uncertainty, the appropriate long-term discount rate to analyze the costs and benefits of long-term investments (e.g., investments to avoid global warming) is lower than that inferred from the short- and mid-term rates. Our results underline, in a simple parsimonious framework, that we should expect a separation between short and long rates. As in Weitzman's studies, the reason for such a separation in our model is the behavior of the economy in the far-left tail, i.e., in the worst-case scenarios. However, the long rate in our framework can be either higher or lower than the short rate, whereas in Weitzman's studies it is always lower because of parameter uncertainty.

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. Alternatively, a lower bound on consumption can be interpreted as subsistence farming or consumption. 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.

The importance of the low consumption 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, e.g., introduced 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 equity premia.[3] Thus, we are not the first to consider the effect of extreme events in consumption based asset pricing. All of these approaches consider the effect of catastrophic risk, either actual or suspected, on agents' valuation for risky bets. However, as we have argued above, it is difficult to envisage economic environments in which agents cannot guarantee themselves at least a subsistence level of consumption. By this reasoning, the equity premium puzzle cannot be resolved by appealing to low probability catastrophic states, because agents always have resources with which they can guarantee themselves at least subsistence consumption.

The paper is organized as follows. In the next section, we introduce the model. In Section 3, we study the term structure and the utility of the representative investor, and finally, Section 4 concludes. All proofs are deferred to an appendix.

2 Model

Consider an economy that evolves between times 0 and T (where T can be finite or infinite), in which there are two sources of the consumption good. The first pays a riskless dividend, $B(t)\,dt$, which grows at a constant rate k over time, so the time $t > 0$ dividend is $B(t) = B_0\, e^{kt}$. We call this asset a bond. For simplicity, we focus on the case where $k = 0$, so $B(t)$ is constant, but our results readily generalize to nonzero growth rates.

The second, 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,

[3] An interesting and unresolved question in the catastrophic events literature is if and how the expected utility of a representative agent who faces decimation should be defined. This "survival problem" has been examined by Majumdar and Hashimzade (2006).

ω is a standard Brownian motion, which generates a standard filtration, $\mathcal{F}_t$, on $t \in [0, T)$. We also define $\hat{\mu} = \mu + \sigma^2/2$.

It will be useful to consider the share of the risky asset in the overall economy and so we define the risky share, $z(t) = D(t)/(B(t) + D(t))$. Notice that if z is constrained to be one, all resources are in the risky asset and the economy collapses to a standard one tree model in line with that in Lucas (1978).[4] It will sometimes be convenient to use $d = \log(z) - \log(1 - z) = \log(D/B)$.

There is a price-taking representative investor with constant relative risk-averse (CRRA) utility and risk-aversion coefficient $\gamma > 0$, who consumes the output of both trees:

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

Here,

$$u(c) = \begin{cases} \log(c), & \gamma = 1, \\ \frac{c^{1-\gamma}}{1-\gamma}, & \gamma \neq 1. \end{cases} \tag{2}$$

The prices of all assets are determined by her valuations, and a standard argument implies that, in equilibrium, an asset that pays out $\xi(t)$, where ξ is an $\mathcal{F}_t$ adapted process satisfying standard conditions, commands an initial price of

$$P_0^T = \frac{1}{u'(B_0 + D_0)} E_0 \left[\int_0^T e^{-\rho s} u'(B(s) + D(s)) \xi(s)\, ds \right]. \tag{3}$$

Equation (3) is the Euler equation that relates the agent's aggregate consumption, her marginal utility and her valuation for all securities, including risk-free ones. When we analyze welfare and utility in Subsection 3.1 below, we establish that prices of risk-free bonds can be expressed as a function of the fraction of the representative agent's wealth held in the risky asset. Therefore, we write bond prices as $P_0^\tau(z)$, where $z = D_0/(B_0 + D_0)$.

A zero-coupon risk-free bond with maturity date τ has the price

$$P_0^\tau = \frac{1}{u'(B_0 + D_0)} E_0 \left[e^{-\rho\tau} u'(B(s) + D(s)) \right]. \tag{4}$$

The τ-period spot rate at time 0 is defined as

$$r_\tau = -\frac{\log(P_0^\tau)}{\tau},$$

[4]The Fisherian consumption model presented in Lucas (1978) follows earlier equilibrium models such as Rubinstein (1976).

while the short rate is defined as

$$r^s = \lim_{\tau \searrow 0} r_\tau,$$

and the long rate, in an economy with $T = \infty$, is defined as

$$r^l = \lim_{\tau \to \infty} r_\tau.$$

Standard stochastic calculus implies dynamics for the capital in the risky sector and for the risky share:

$$\begin{aligned} dD &= D\left(\widehat{\mu}\, dt + \sigma\, d\omega\right), \\ dz &= z(1-z)\left(\widehat{\mu}\, dt + \sigma\, d\omega\right) - z^2(1-z)\sigma^2\, dt. \end{aligned} \tag{5}$$

We note that when $\mu > 0$, 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.[5] In what follows, we focus our attention on the economically interesting case of $\mu > 0$.

Cochrane, Longstaff, and Santa-Clara (2008) also characterize the economy in terms of the relative share of each asset, z, and express dynamics for the share. However, they assume that both trees are risky with the same growth rate in most of their analysis. In our formulation, the difference between the drifts of the two trees is μ, whereas in their case the difference is zero. This leads to different formulae for the dynamics of the risky share (in our model, Equation (5)).

Through the Feynman-Kac equations, a general partial differential equation (PDE) for the price of a zero coupon bond can be derived:

Proposition 1 *The price, at time s, of a τ-maturity zero coupon bond, where $s + \tau \leq T$, is $P_s^\tau(z) = F(s,z)$, where F is the solution to the following PDE:*

$$\begin{aligned} F_t + \frac{1}{2}\sigma^2 z^2 (1-z)^2 F_{zz} + (\gamma\widehat{\mu} z(1-z) - \sigma^2(1+\gamma)z^2(1-z))F_z \\ - \left(\rho + \gamma\widehat{\mu} z - \frac{1}{2}\gamma(\gamma+1)\sigma^2 z^2\right) F = 0. \end{aligned} \tag{6}$$

[5] The convergence also holds a.s. for $\mu \neq 0$, but not for $\mu = 0$.

$$F(z, s+\tau) \equiv 1, \tag{7}$$

$$s \le t \le \tau,$$
$$0 \le z \le 1.$$

We note that no boundary conditions are needed at the $z = 0$ and $z = 1$ boundaries. The technical reason is that both the first order (F_z) and the second order (F_{zz}) terms vanish at these boundaries, so that the PDE turns into a first order ODE in t. The boundary conditions in the two-tree model are further analyzed in a companion paper, Parlour, Stanton, and Walden (2009).

Intuitively, boundary conditions are not needed because both boundaries are absorbing: If the economy is in either the risk-free ($z = 0$) or the full-risk ($z = 1$) state, it stays in this state as there is no way to transfer capital between the trees. At either of these end points, the model collapses to the one-tree framework and bond prices are deterministic. Consider Equation (6) evaluated at $z = 0$ and $z = 1$ respectively. The solution to $F_t - \left(\rho + \gamma\widehat{\mu}z - \frac{1}{2}\gamma(\gamma+1)\sigma^2 z^2\right) F = 0$, $F(s+\tau) = 1$ is $P_s^\tau = F(s) = e^{-\rho\tau}$ when $z = 0$ and is the discount rate in a deterministic economy. Similarly, when $z = 1$ the economy is the standard CRRA-lognormal framework and the price is $P_s^\tau = F(s) = e^{-(\rho+\gamma\widehat{\mu}-\gamma(\gamma+1)\sigma^2/2)\tau}$.

3 Characterization of the economy

The normalized expected utility of the representative agent, or the welfare, is helpful in understanding the term structure as the valuation of any risk-free bond depends on the agent's expected marginal utility. We first characterize the welfare and then proceed to characterize the term structure.

First, observe that without loss of generality, we can normalize the size of the economy to 1. That is, we assume that $B_0 + D_0 = 1$. The homogeneity of the utility function implies that such a normalization has no impact on the term structure, and that the expected utility, when $\gamma \neq 1$, is scalable as $U(t, z, B+D) = (B+D)^{1-\gamma}U(t, z, 1) \stackrel{\text{def}}{=} (B+D)^{1-\gamma}w(t, z)$, where $w(t, z)$ is the normalized welfare at share z, $w(t, z) \stackrel{\text{def}}{=} U(t, z, 1)$. In the special case of log-utility, i.e., when $\gamma = 1$, we have $U(t, z, B+D) = \frac{\log(B+D)\left(1-e^{-\rho(T-t)}\right)}{\rho} + w(t, z)$. In the infinite horizon case, the welfare is not time dependent, $w = w(z)$.

3.1 Expected utility

The expected utility of the representative investor can be found by solving a PDE. In the appendix, we develop the PDE that characterizes the representative agent's welfare and state this as Proposition 6. Here, we report the infinite horizon case, which can be calculated directly.[6] We have

Proposition 2 *In the infinite horizon economy, $T = \infty$, define $q = \sqrt{\mu^2 + 2\rho\sigma^2}$. Suppose that*
(i) $\gamma = 1$. Then, if the initial risky share is $0 < z < 1$, the expected utility of the representative agent is

$$\begin{aligned} w(z) &= \frac{1}{2\rho}\Big(\left(2\mu^2 + \sigma^2(2\rho + q) + \mu(\sigma^2 + 2q)\right) {}_2F_1\left(1, \frac{q-\mu}{\sigma^2}, \frac{q-\mu}{\sigma^2} + 1, \frac{z-1}{z}\right) \\ &+ 2\frac{z}{z-1}\left(\mu^2 + \rho\sigma^2 - \mu q\right) {}_2F_1\left(1, \frac{q+\mu}{\sigma^2} + 1, \frac{q+\mu}{\sigma^2} + 2, \frac{z}{z-1}\right)\Big) \\ &/ \left(\mu^2 - \mu q + 2\rho(\sigma^2 + q)\right), \end{aligned}$$

where ${}_2F_1$ is the hypergeometric function. Also, $w(0) = 0$ and $w(1) = \frac{\mu}{\rho^2}$.
(ii) If $\gamma > 1$: then if the initial risky share is $0 < z < 1$, the expected utility of the representative agent is

$$\begin{aligned} w(z) &= \frac{(1-z)^{1-\gamma}}{q(1-\gamma)} \\ &\times \Bigg[\left(\frac{1-z}{z}\right)^{\frac{\mu-q}{\sigma^2}} \left(V\left(\frac{1-z}{z}, \gamma + \frac{q-\mu}{\sigma^2}, 1-\gamma\right) + V\left(\frac{1-z}{z}, \gamma + \frac{q-\mu}{\sigma^2} - 1, 1-\gamma\right)\right) \\ &+ \left(\frac{z}{1-z}\right)^{-\frac{q+\mu}{\sigma^2}} \left(V\left(\frac{z}{1-z}, \frac{q+\mu}{\sigma^2}, 1-\gamma\right) + V\left(\frac{z}{1-z}, \frac{q+\mu}{\sigma^2} + 1, 1-\gamma\right)\right)\Bigg]. \end{aligned}$$

Here, $V(y, a, b) \stackrel{\text{def}}{=} \int_0^y t^{a-1}(1+t)^{b-1}dt$ is defined for $a > 0$. Also, $w(0) = \frac{1}{\rho(1-\gamma)}$. Moreover, define $x \stackrel{\text{def}}{=} \rho + (\gamma-1)\mu - (\gamma-1)^2\frac{\sigma^2}{2}$. Then, if $x > 0$, $w(1) = -\frac{1}{x}$. If, on the other hand, $x \leq 0$, 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 x, so we prefer using the function V, which is real valued. Also, since the Beta function and the hypergeometric function satisfy

[6] Here, we treat the infinite horizon case, $T = \infty$, as the limit as $T \to \infty$.

the relationship, $B(x,a,b) = {}_2F_1(1-b,a,a+1,x)$, the cases $\gamma = 1$ and $\gamma > 1$ have similar functional forms.

Part (ii) of Proposition 2 provides intuition for why the risk-free tree alters the economy relative to the one-tree model. In it, we define $x = \rho + (\gamma-1)\mu - (\gamma-1)^2\frac{\sigma^2}{2}$. Suppose that $x \leq 0$, then expected utility at $z = 1$ is infinitely negative, conversely for $x > 0$, the utility is strictly positive. The case of $z = 1$ corresponds to the standard Lucas-Tree model, with one risky tree, in which bonds are in zero net supply. However, for z arbitrarily close to but less than one, expected utility is finite because of the insurance provided by the risk-free tree in bad states of the world. Thus, even in an economy with a very small risk-free tree, where real variables behave much like in the one-tree model, the representative agent's asset valuations differ markedly from those in the one-tree model.

To understand the importance of the insurance offered by the risk-free asset, we analyze the point at which $x = 0$, which we refer to as the *break-point* of the model. In particular, we consider its implications for risk aversion. Consider the following rough parametrization: Suppose that the consumption growth rate is $\mu = 1\%$ and that the consumption volatility is $\sigma = 4\%$. 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%, so the number is off by a factor less than two. We make the upward adjustment to reflect the arguments in both Parker (2001) and Gabaix and Laibson (2001) that consumption adjustment costs may artificially reduce measured consumption volatility. We take the personal discount rate to be $\rho = 1\%$. Given these parameters, if $x = 0$, then the risk-aversion break-point coefficient is

$$\gamma = 1 + \frac{q+\mu}{\sigma^2}, \tag{8}$$

where $q = \sqrt{\mu^2 + 2\rho\sigma^2}$, as defined in Proposition 2.

The above parameter values yield a risk aversion value of $\gamma = 14.43$. Thus, if $\gamma > 14.43$ — a high, but not extreme number — expected utility is negatively infinite. If consumption volatility is halved, so that $\sigma = 2\%$ then the break-point occurs at $\gamma = 52.0$. Now, even if γ is lower than this, but still close to the break-point, the behavior of the representative investor will still be very sensitive to z close to $z = 1$.

Even when $x > 0$, the differences between the one-tree and two-tree models may be large. Figure 1 illustrates the expected utility as a function of z, the share in the risky asset, for the case of log-utility, $\gamma = 1$. We see that the utility function is quite flat in the interior, but decreases sharply close to the boundaries $z = 0$ and $z = 1$, suggesting that these (one-tree) boundary cases may be quite different from the interior case, even though it is always the case that $x > 0$ when $\gamma = 1$.

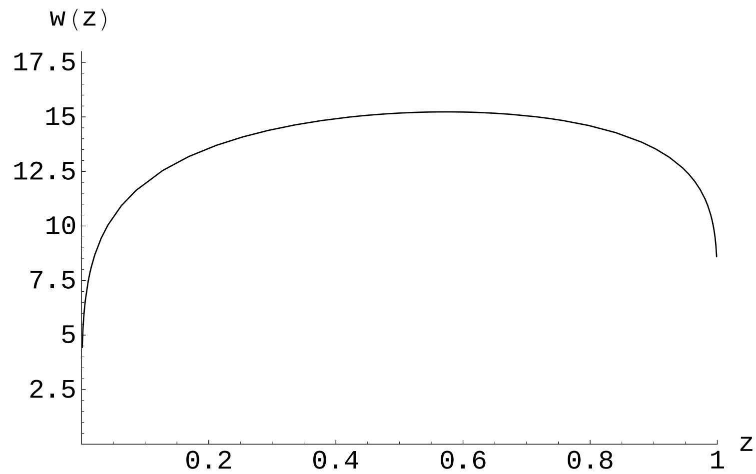

Figure 1: **Expected utility as a function of z, the risky asset share. Parameters:** $\rho = 1/8$, $\sigma = 2$, $\mu = 1/10$, $\gamma = 1$. $w(0) = 0$ **and** $w(1) = 6.4$.

Also, when $T < \infty$, expected utility will be finite for arbitrary z, regardless of T. However, if the time horizon is large, we should again expect the representative agent to be sensitive to z changes close to $z = 1$, similar to the case in Figure 1 with $T = \infty$ and $\gamma = 1$. Thus, if the equilibrium is close to or beyond the break point defined in Equation (8), caution should be taken with quantitative implications of the one-tree model. With this in mind, we move on to our main objective, studying the term structure of interest rates.

3.2 The Term Structure

We first recall some well-known properties of the one-tree model, which formally occurs as a special case of our two-tree model when $z = 1$. For $z = 1$ the yield curve is constant,

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

The expected return of stocks is

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

implying that the equity premium is

$$g - r = \gamma\sigma^2.$$

With an equity premium of 5% and a volatility of $\sigma = 4\%$, this implies a risk-aversion parameter of $\gamma = 31$. The sensitivity of the interest rate to γ is shown in Figure 2, for $\sigma = 1\%$, $\sigma = 2\%$ and $\sigma = 4\%$.

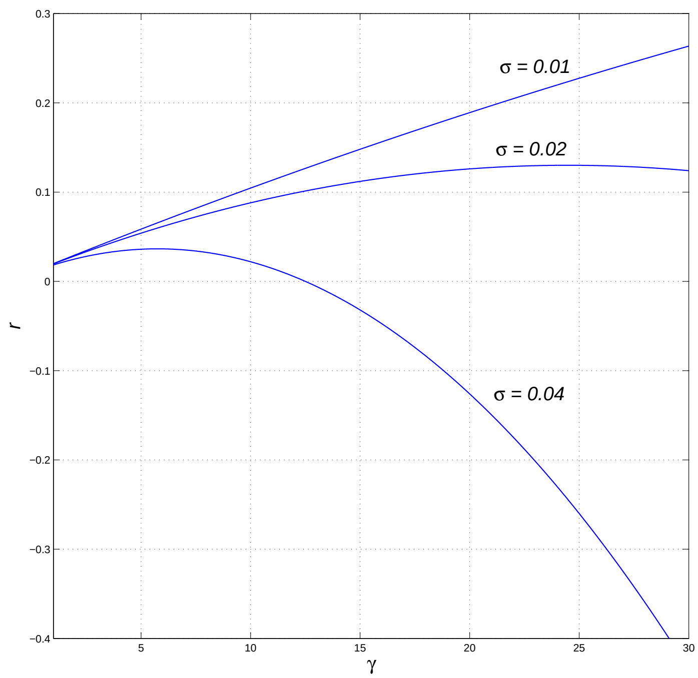

Figure 2: **Risk-free rate in one-tree model. Parameters:** $\rho = 1\%$**,** $\mu = 1\%$**,** $\gamma = 31$**,** $\sigma = 1\%, 2\%, 4\%$**.**

The main implication of Figure 2 is that for a high γ (in this case $\gamma = 31$), which is needed to match the equity premium, the interest rate is very sensitive to parameter values, i.e., to the combination of ρ, μ and σ. This is the risk-free rate puzzle. For $\gamma = 30$, when $\sigma = 0.01$, the risk-free rate is 26%, whereas when $\sigma = 0.04$, it is -43%. Thus, as for example noted in Campbell (2003), a reasonable interest rate only occurs as a "knife-edge" case.

In the two-tree model, the term structure is no longer constant. We have:

Proposition 3 *Consider an infinite horizon economy in which the relative size of the sec-*

tors is $d = \log(z/(1-z)) = \log(D/B)$. Then, the price of a τ-period bond is given by

$$P^\tau = (1+e^d)^\gamma e^{-\rho\tau} \frac{1}{\sqrt{2\pi\sigma^2\tau}} \int_{-\infty}^{\infty} \frac{e^{-(y-d-\mu\tau)^2/(2\sigma^2\tau)}}{(1+e^y)^\gamma} dy. \tag{9}$$

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 (9) can be expressed in the following form:

(i) If $\gamma = 1$:

$$\begin{aligned} P^\tau &= \frac{(1+e^d)^\gamma e^{-\rho\tau-(d+\mu\tau)^2/(2\sigma^2\tau)}}{2} \times \\ & \sum_{n=0}^{\infty} (-1)^n a_n \left(F\left(\frac{d+\mu\tau+n\tau\sigma^2}{\sqrt{2\sigma^2\tau}}\right) + F\left(\frac{-d-\mu\tau+(n+\gamma)\tau\sigma^2}{\sqrt{2\sigma^2\tau}}\right)\right) \end{aligned} \quad , (10)$$

(ii) If $\gamma \geq 1$:

$$\begin{aligned} P^\tau = & \frac{(1+e^d)^\gamma e^{-\rho\tau-(d+\mu\tau)^2/(2\sigma^2\tau)}}{2} \times \lim_{\epsilon\searrow 0} \\ & \sum_{n=0}^{\infty} (-1)^n e^{-\epsilon n} a_n \left(F\left(\frac{\epsilon+d+\mu\tau+n\tau\sigma^2}{\sqrt{2\sigma^2\tau}}\right) + F\left(\frac{\epsilon-d-\mu\tau+(n+\gamma)\tau\sigma^2}{\sqrt{2\sigma^2\tau}}\right)\right) \end{aligned} , (11)$$

(iii) If $\gamma = 1$ and $\frac{d+\mu\tau}{\sigma^2\tau} = m \in \mathbb{N}$, then

$$P^\tau = \frac{(1+e^d)e^{-\rho\tau-m^2\sigma^2\tau/2}}{2}\left(1 + 2\sum_{n=1}^{m-1}(-1)^n e^{n^2\sigma^2\tau/2}\right). \tag{12}$$

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.

We note that Equations (10) and (12) are special cases of Equation (11).[7] 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 his characterization is less specific than those in Proposition 3.

We use Equations (11)–(12) to study the yield curve with parameters $\rho = 1\%$, $\sigma = 20\%$,

[7] The reason why the limit as $\epsilon \searrow 0$ is needed when $\gamma > 1$ is that the series diverges otherwise. This is expanded upon in the proof in the appendix.

$\mu = 4\%$, $z = 90\%$, and risk-aversion coefficients between 2 and 5. The results are shown in Figure 3. The choice of $z = 90\%$ at $t = 0$ means that the risky tree initially dominates

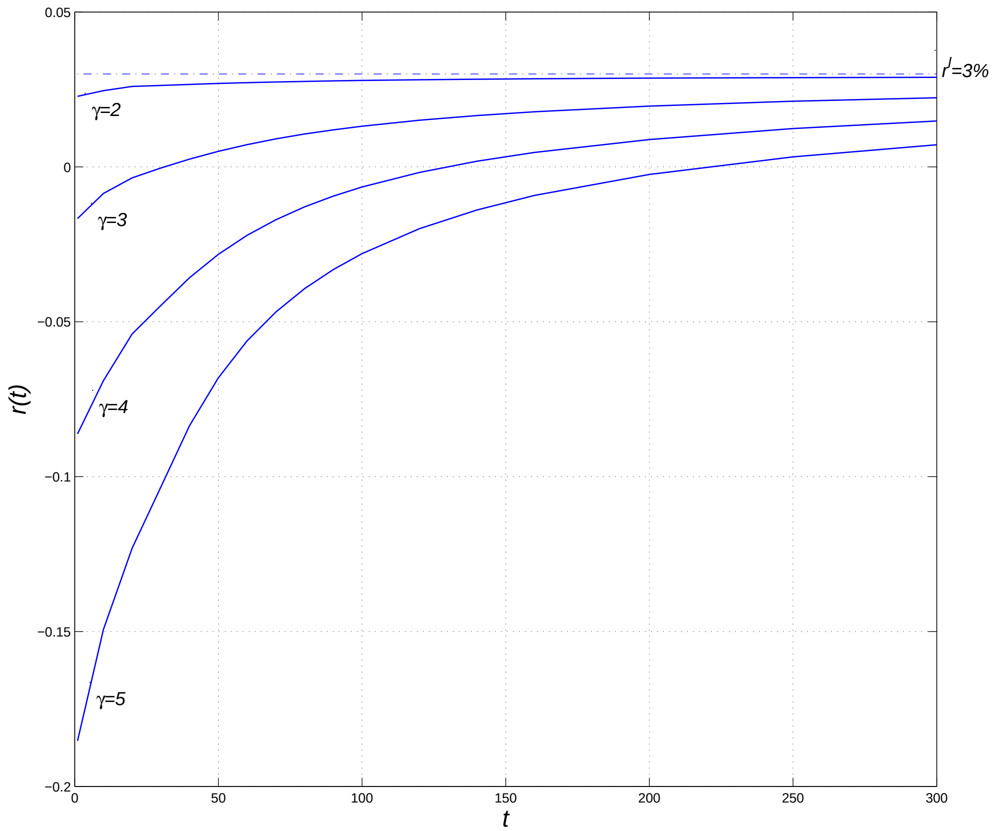

Figure 3: **Term structure of interest rates in the two-tree model. Parameters: $\rho = 1\%$, $\sigma = 20\%$, $\mu = 4\%$, $z = 90\%$ $\gamma = 2, 3, 4, 5$.**

the economy and, since $\hat{\mu} > \frac{\sigma^2}{2}$, 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 increases 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 4 *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

$$r^l = \rho + \frac{1}{2} \times \frac{\mu^2}{\sigma^2}. \tag{13}$$

If, on the other hand, $z \in (0,1)$ and $\mu > \gamma\sigma^2$, the long-term rate is

$$r^l = \rho + \gamma \left(\mu + \frac{\sigma^2}{2}\right) - \gamma(\gamma+1)\frac{\sigma^2}{2}. \tag{14}$$

Thus, the short rate has the same structure as in the one-tree model and, as long as $\mu > \gamma\sigma^2$, the long rate is the same as the long rate in the economy with only one risky tree. 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 $\gamma\sigma^2$ is the equilibrium equity premium, which is approximately 5–8% per year,[8] while μ, the consumption growth rate, is about 1% per year, $\mu < \gamma\sigma^2$; this suggests that the long-term rate is γ (risk-aversion) independent in practice.[9]

This γ-independence stands in stark contrast to the results in the one-tree model, where, as we have seen, the interest rate is very sensitive to risk aversion.[10] Specifically, in the two-tree model, the long rate is always greater than the personal discount rate, $r_l > \rho$, regardless of the aggregate risk aversion in the economy. In our previous example, with $\mu = 1\%$, $\sigma = 4\%$ and $\rho = 1\%$, the long-term rate is $r^l = 4.125\%$, regardless of γ and also of $z \in (0,1)$ in the two-tree model, whereas it is $r^l = \rho + \gamma\widehat{\mu} - \gamma(\gamma+1)\sigma^2/2 = -45\%$ in the one-tree model.

Notice that this γ-independence offers a resolution to the risk-free rate puzzle at the long end of the term structure. The insurance offered by the (in the long run vanishingly

[8] See, e.g., the discussion in Brealey, Myers, and Allen (2008), page 180.

[9] 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.

[10] We have verified that the formula is indeed correct by numerically integrating Equation (4) 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 (13), since $3.5\% < 2.5 \times 20\%^2$). On the contrary, Equation (14) 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_0, D_0 and γ, it is easily verified that r^l does not depend on these parameters.

small) risk-free tree is thus extremely important for long-term discount rates. Technically, the risk-free tree bounds marginal utility from above in bad states of the world, which has a major impact on what the investor is willing to pay for risk-free assets. Therefore, even if risk aversion is high, the discount rate will be modest. In this sense, this approach is opposite of Rietz (1988) and Barro (2005), who explain the equity premium puzzle by introducing rare catastrophic events, while keeping γ low. Our approach on the other hand, which puts a limit on aggregate losses in the economy, allows for a higher γ without affecting the long-term rate, which leads to higher equity premia. We stress, though, that the main point of our analysis is to better understand the determinants of long rates, not to solve the equity premium puzzle. Further, as we alluded to in the introduction, a subsistence level of consumption seems to be the most plausible assumption.

The differences between the long rates in the two models further underlines the fragility of the CRRA-lognormal model in longer time horizons, when risk aversion is high. Regardless of how close z is to 1 in the two-tree model, the long-term rate is drastically different from the case when z is identically equal to 1. Similar to the expected utility results, the differences between the two models are driven by the insurance the risk-free tree provides in the far-left tails.

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. 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.

We first clarify the effect of the risk-free sector on the long-term rate.

Corollary 1 *(i) Regardless of the size of the risk-free sector, for a fixed choice of γ, ρ, σ and μ, then the existence of such a sector means that the long-term rate is greater than or equal to the long-term rate without such a sector.*
(ii) If $\mu < \gamma\sigma^2$, then the long-term rate in an economy with a risk-free sector is strictly higher than in an economy without.

Since the short rate is equal to the long rate in the one-tree economy, and the short rates are basically the same in the two-tree and one-tree economies when z is close to 1, this also implies the following result:

Corollary 2 *If $\mu < \gamma\sigma^2$, then for z close to 1, $r^l > r^s$ in the two-tree economy.*

Thus, for z close to 1, our results are the opposite of those obtained by Weitzman: The short rate is lower than the long-term rate. When z is close to 1, the consumption growth rate and volatility are stable; they depend on z, which is almost constant, as seen from Equation (5), and the fact that $z \to_p 1$. However, when z is far away from 1, the consumption growth rate and volatilities will change over time as z converges to 1. Specifically, the consumption growth rate will increase from $z\mu$ to μ. Thus,

Corollary 3 *(i) In an economy in which the consumption growth rate is expected to be stable over time, $r^l > r^s$.*
(ii) Whereas, if $r^s \geq r^l$ then the consumption growth rate is expected to increase.
(iii) In the long run, the yield curve will on average be upward sloping.

This corollary is important in evaluating the appropriate long term discount rates across different countries. Specifically, if we assume that there are sufficient barriers between countries, then in mature economies such as the United States, consumption growth is stable and the appropriate discount rate to use should be larger than the shorter, observed market rates. By contrast, in rapidly developing nations such as India, where the consumption growth rate is expected to increase, calculating the appropriate discount rate from market rates is somewhat more complex. There will be long time periods in which the appropriate rate at which to discount long run projects will be lower than the observed shorter, market rates.

We have not yet shown that there are states in the world in which $r^l < r^s$. In Figure 4 we show the yield curve for different values of z between $z = 0.1$ and $z = 0.5$, for $\mu = 1/3$, $\sigma = 1$, $\rho = 0$, $\gamma = 1$. We see that there are indeed regions in which the yield curve is downward sloping. A further analysis leads to the following conditions that determine the sign of $r^l - r^s$:

Proposition 5 *Define*

$$q_2 \stackrel{\text{def}}{=} \sqrt{\left(\frac{\mu}{\sigma^2} + \frac{1}{2}\right)^2 - \frac{1+\gamma}{\gamma} \times \frac{\mu^2}{\sigma^4}}.$$

Then, if $\mu < \gamma\sigma^2$, for $z \in (0, 1)$,
(i) The term structure slopes up, $r_l > r_s$, if $z < \frac{1}{1+\gamma}\left(\frac{\mu}{\sigma^2} + \frac{1}{2} - q_2\right)$ or $z > \frac{1}{1+\gamma}\left(\frac{\mu}{\sigma^2} + \frac{1}{2} + q_2\right)$,
(ii) The term structure slopes down, $r_l < r_s$, if $\frac{1}{1+\gamma}\left(\frac{\mu}{\sigma^2} + \frac{1}{2} - q_2\right) < z < \frac{1}{1+\gamma}\left(\frac{\mu}{\sigma^2} + \frac{1}{2} + q_2\right)$.

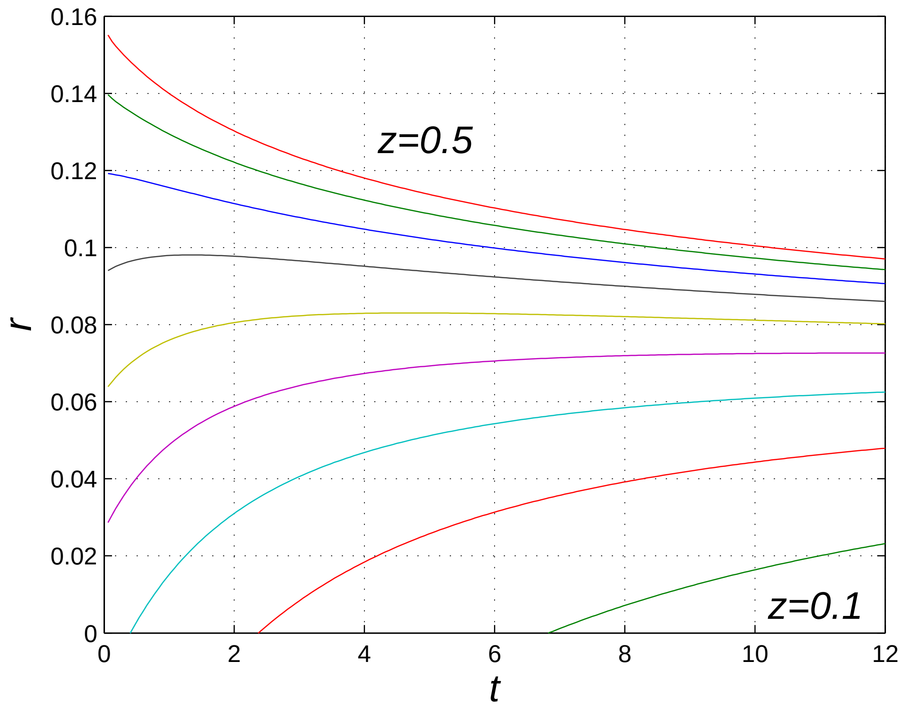

Figure 4: **Zero-coupon yield curve 0-12 years, for varying z. Parameters:** $\mu = 1/3$, $\sigma^2 = 1$, $\rho = 0$, $\gamma = 1$.

(iii) $r_l - r_s$ *is decreasing in* z *if* $z < \frac{1}{1+\gamma}\left(\frac{1}{2} + \frac{\mu}{\sigma^2}\right)$.
(iv) $r_l - r_s$ *is increasing in* z *if* $z > \frac{1}{1+\gamma}\left(\frac{1}{2} + \frac{\mu}{\sigma^2}\right)$.

What is the economic intuition behind these results? Suppose that the economy is far away from the long-term z, which is close to 1. In this case, the instantaneous short rate is just the representative agent's expected relative change in marginal utility. Market clearing that is imposed on the economy as it evolves means that the agent must be indifferent between consuming today and transferring consumption to tomorrow. Suppose that he expects a lower marginal utility next period. He therefore values consumption next period less (a decrease in demand for consumption tomorrow). Market clearing imposes that interest rates rise. Therefore, the term structure slopes up.

There is a large empirical literature that examines the relationship between the term structure and the macro economy (Ang, Piazzesi, and Wei (2006) present a survey). In this literature, a stylized fact is that the higher the slope of the term structure, the larger expected GDP growth.[11] The relationship between the slope of the term structure and the

[11] Kessel (1965); Harvey (1986, 1988, 1989, 1993); Laurent (1988); Stock and Watson (1989); Chen (1991);

growth rate is consistent with our model. Indeed, part (iv) of Proposition 5 provides a condition under which the slope of the term structure is increasing in z (above a threshold value for z). When the economy is naturally in this region, then a larger slope implies that the growth rate will be higher. There is evidence (Bernard and Gerlach (1998) and Bonser-Neal and Morley (1997)) that the term structure slope has different predictive power in different countries. In as much as different parameters govern the evolution of different economies, our model suggests that there should be international differences in the predictive power of the term structure.

Other features of the observed term structure are consistent with those generated by the two-tree framework. Inspection of Figure 4 indicates that the term structure can be hump-shaped for some values of z. The presence of a hump is one of the stylized properties of the real world term structure (see Nelson and Siegel (1987)). The curvature, however, is quite small and is even smaller for lower values of σ^2. Ang, Bekaert, and Wei (2007) analyze ten years of TIPS data to characterize the real term structure in the United States. They find that, unconditionally, the term structure of real rates is somewhat flat, with a peak at the one year maturity, and sometimes slopes downward. They find that the short end is quite variable and on average there is little or no term spread.

Thus, in addition to pointing out the sensitivity of the term structure in CRRA-lognormal framework, our model is in line with several stylized facts documented in the empirical term structure literature, and leads to some novel predictions. It is promising that such a rich structure is provided by such a simple parsimonious model.

4 Concluding remarks

Finding the appropriate rate at which to discount cashflows is fundamental to finance. As society contemplates investments with long horizons, it is important for us to be able to address basic questions about how to value cashflows in the distant future. We have shown that if the representative agent has access to a minimum subsistence level of consumption, then her valuation of long-run cashflows is markedly different from her valuation in the absence of such a guaranteed minimum. Indeed, at longer horizons and with highly risk averse agents, the insurance offered by the additional risk-free sector changes the representative investor's valuation of risk-free investments and makes the long rate completely independent of the agent's risk aversion. Our results are consistent with the view that long-term discount rates are determined by factors different from those that drive short- and mid-term discount

Estrella and Hardouvelis (1991); Estrella and Mishkin (1998); Dotsey (1998); Hamilton and Kim (2002); Moody and Taylor (2003) consider the relationship between GDP growth and term spreads in the US, while Jorion and Mishkin (1991); Harvey (1991); Estrella and Mishkin (1997); Plosser and Rouwenhorst (1994); Bernard and Gerlach (1998) consider the international evidence.

rates, giving us greater flexibility in resolving the risk-free rate puzzle and providing insight into how to discount long-term cashflows.

The fact that different variables drive the long and short rates in our economy means that, unlike the flat term structure in a one-tree economy, the term structure in a two-tree economy can exhibit more of the shapes observed in practice. For example, the term structure frequently slopes upwards, and can be hump shaped. Further, the slope of the term structure in a two-tree economy conveys information about expected future consumption growth rates, consistent with empirical results such as those in Ang, Piazzesi, and Wei (2006). This suggests that the model has appropriate qualitative features.

More generally, our results illustrate that the standard long-horizon Lucas model with a CRRA representative investor and a log-normal consumption process is highly sensitive to small perturbations, especially when risk aversion is high. One test of how severe this sensitivity may be is given by checking whether the parameters of the model imply that the equilibrium is close to, or even above, a break point. Our framework also suggests that there are non-monotonicities in the relationship between the term structure and future economic aggregates, which future empirical research should account for.

Welfare in finite horizon case

Proposition 6 *The welfare, w, at time $t < T$ is, the solution to the following PDE:*

$$
\begin{aligned}
w_t + \frac{1}{2}\sigma^2 z^2 (1-z)^2 w_{zz} + \left[\widehat{\mu} z(1-z) - \sigma^2 \gamma z (1-z)^2\right] w_z \\
- \left[\rho + \widehat{\mu}(\gamma - 1)(1-z) - \frac{1}{2}\sigma^2 \gamma(\gamma-1)(1-z)^2\right] w + R_\gamma(z) = 0,
\end{aligned}
$$

where

$$
R_\gamma(z) = \begin{cases} \frac{1-e^{-\rho(T-t)}}{\rho}\left(\widehat{\mu} z - \frac{\sigma^2 z^2}{2}\right), & \gamma = 1, \\ \frac{1}{1-\gamma}, & \gamma \neq 1. \end{cases}
$$

The boundary condition is

$$
w(z, T) = 0. \tag{15}
$$

Proof of Proposition 6:

Using Equation (1) and the Feynman-Kac equations, we get that $U = w(0, z)$, where $w(t, z)$ solves the partial differential equation

$$
\begin{aligned}
w_t + \frac{1}{2}\sigma^2 z^2 (1-z)^2 w_{zz} + \left[\widehat{\mu} z(1-z) - \sigma^2 \gamma z (1-z)^2\right] w_z \\
- \left[\rho + \widehat{\mu}(\gamma - 1)(1-z) - \frac{1}{2}\sigma^2 \gamma(\gamma-1)(1-z)^2\right] w + R_\gamma(z) = 0,
\end{aligned}
$$

where

$$
R_\gamma(z) = \begin{cases} \frac{1-e^{-\rho(T-t)}}{\rho}\left(\widehat{\mu} z - \frac{\sigma^2 z^2}{2}\right), & \gamma = 1, \\ \frac{1}{1-\gamma}, & \gamma \neq 1. \end{cases}
$$

The boundary condition is

$$
w(z, T) = 0.
$$

A similar argument to the proof of Proposition 1 shows that no boundary conditions are needed at the $z = 0$ and $z = 1$ boundaries. ■

Proofs

Proof of Proposition 1:

Using Equation (4), that $u'(c) = c^{-\gamma}$, and the Feynman Kac formula, we obtain the following pricing equation for any asset P:

$$P_t + \frac{1}{2}\sigma^2 D^2 P_{DD} + \left[\widehat{\mu} D - \frac{\sigma^2 \gamma D^2}{B+D}\right] P_D \\ - \left(\rho + \widehat{\mu}\gamma \frac{D}{B+D} - \frac{1}{2}\sigma^2\gamma(\gamma+1)\frac{D^2}{(B+D)^2}\right) P + \delta(B, D, t) = 0. \quad (16)$$

Using $z = D/(B+D)$, for bonds this equation can be simplified to

$$p_t + \frac{1}{2}\sigma^2 z^2 (1-z)^2 p_{zz} + \left[\gamma\widehat{\mu} z(1-z) - \sigma^2(1+\gamma)z^2(1-z)\right] p_z \\ - \left[\rho + \gamma\widehat{\mu} z - \frac{1}{2}\gamma(1+\gamma)\sigma^2 z^2\right] p = 0, \quad (17)$$

with terminal condition

$$F(z, s+\tau) \equiv 1,$$

To show that no boundary conditions are needed at the $z = 0$ and $z = 1$, we show that the PDE without boundary conditions is well posed (Egorov and Shubin (1992)). The concept of well-posedness additionally requires the solution to depend continuously on initial and given boundary conditions, and ensures that the given boundary conditions are enough.

We note that the PDE (17) is parabolic in the interior, both the parabolic and hyperbolic terms vanish at the boundaries. This can be viewed as a special case of an outflow boundary for a hyperbolic PDE. For outflow boundaries to hyperbolic equations, no boundary conditions are needed, i.e., if the Cauchy problem is well posed, then the initial-boundary value with an outflow boundary is well-posed without a boundary condition (Kreiss and Lorenz (1989)), and it suggests that no boundary conditions are needed.

To show that this is indeed the case, we use the energy method to show that the operator $Pw \stackrel{\text{def}}{=} -\left[\rho + \gamma\widehat{\mu} z - \frac{1}{2}\gamma(1+\gamma)\sigma^2 z^2\right] w + (z(1-z)\gamma\hat{\mu} - (1+\gamma)z^2(1-z)\sigma^2)w_z + \frac{\sigma^2}{2} z^2(1-z)^2 w_{zz} \stackrel{\text{def}}{=} aw + bw_z + cw_{zz}$ is maximally semi-bounded, i.e., we use the L_2 inner product $\langle f, g\rangle = \int_0^1 f(x)g(x)dx$, and the norm $\|w\|^2 = \langle w, w\rangle$, and show that for any smooth function, w, $\langle w, Pw\rangle \le \alpha\|w\|^2$, for some $\alpha > 0$.[12] This allows us to bound the growth of $\frac{d}{d\tau}\|w(t,\cdot)\|^2$ by $\frac{d}{d\tau}\|w(t,\cdot)\|^2 \le \alpha\|w\|^2$, since $\frac{1}{2} \times \frac{d}{d\tau}\|w(t,\cdot)\|^2 = \langle w, Pw\rangle$. Such a growth bound, in turn, ensures well-posedness (see Kreiss and Lorenz (1989) and Gustafsson, Kreiss, and Oliger (1995)).

[12] Since we impose no boundary conditions, it immediately follows that P is maximally semi-bounded if it is semi-bounded.

By integration by parts, defining $q = \max_{z\in[0,1]} |\rho + \gamma\widehat{\mu} z - \frac{1}{2}\gamma(1+\gamma)\sigma^2 z^2|$, we have

$$\begin{aligned}
\langle w, Pw\rangle &= \langle w, aw\rangle + \langle w, cw_z\rangle + \langle w, dw_{zz}\rangle \\
&\leq q\|w\|^2 + \frac{1}{2}\left(\langle w, cw_z\rangle - \langle w_z, cw\rangle - \langle w, c_z w\rangle + [w^2 c]_0^1\right) - \langle w_z, dw_z\rangle - \langle w, d_z w_z\rangle + [wdw_z]_0^1 \\
&= q\|w\|^2 - \langle w, c_z w\rangle - \langle w_z, dw_z\rangle - \langle w, d_z w_z\rangle \\
&\leq \left(q + \sigma^2\right)\|w\|^2 - \frac{\sigma^2}{2}\int_0^1 z^2(1-z)^2 w_z^2 dz.
\end{aligned}$$

Here, the last inequality follows from

$$\begin{aligned}
-\langle w_z, dw_z\rangle - \langle w, d_z w_z\rangle &= \frac{\sigma^2}{2}\int_0^1 z(1-z)\left(-z(1-z)w_z^2 - (2-4z)ww_z\right)dz \\
&\leq \frac{\sigma^2}{2}\int_0^1 z(1-z)\left(-z(1-z)w_z^2 + 2|w||w_z|\right)dz \\
&\leq \frac{\sigma^2}{2}\int_0^1 z(1-z)\left(-z(1-z)w_z^2 + \frac{z(1-z)}{2}w_z^2 + \frac{2}{z(1-z)}w^2\right)dz \\
&= \sigma^2\|w\|^2 - \frac{\sigma^2}{2}\int_0^1 z^2(1-z)^2 w_z^2 dz,
\end{aligned}$$

where we used the relation $|u||v| \leq \frac{1}{2}(\delta|u| + |v|/\delta)$ for all u, v for all $\delta > 0$. We therefore have the estimate

$$\frac{d}{d\tau}\|w\|^2 \leq \left(q + \sigma^2\right)\|w\|^2.$$

We have thus derived an energy estimate for the growth of $\|w\|^2$, and well-posedness follows immediately from the theory in Kreiss and Lorenz (1989) and Gustafsson, Kreiss, and Oliger (1995).

■

Proof of Proposition 2: For $\gamma = 1$,

$$\begin{aligned}
w(z) &= E\left[\int_0^\infty e^{-\rho s}\log\left(B_0\Big(1 + \frac{D_0}{B_0}e^{y(s)}\Big)\right)ds\right] \\
&= \frac{\log(1-z)}{\rho} + E\left(\int_0^\infty e^{-\rho s}\log\left(1 + \frac{z}{1-z}e^{y(s)}\right)ds\right] \\
&= \frac{\log(1-z)}{\rho} + \int_0^\infty\int_{-\infty}^\infty \frac{1}{\sqrt{2\pi\sigma^2 s}}e^{-\rho s}\log\left(1 + \frac{z}{1-z}e^{y(s)}\right)e^{-(y-\mu s)^2/(2\sigma^2 s)}dy\,ds \\
&= \frac{\log(1-z)}{\rho} + \int_{-\infty}^\infty e^{\frac{y\left(\mu - \sqrt{\mu^2 + 2\rho\sigma^2}\,\mathrm{sign}(y)\right)}{\sigma^2}}\frac{\log\left(1 + \frac{z}{1-z}e^y\right)}{\sqrt{\mu^2 + 2\rho\sigma^2}}dy \\
&= G(z; \mu, \sigma^2, \rho),
\end{aligned}$$

where

$$
\begin{aligned}
G(z;\mu,\sigma^2,\rho) &= \Big(\left(2\mu^2+\sigma^2(2\rho+q)+\mu(\sigma^2+2q)\right)\ {}_2F_1\left(1,\frac{q-\mu}{\sigma^2},\frac{q-\mu}{\sigma^2}+1,\frac{z-1}{z}\right) \\
&+ 2\frac{z}{z-1}\left(\mu^2+\rho\sigma^2-\mu q\right)\ {}_2F_1\left(1,\frac{q+\mu}{\sigma^2}+1,\frac{q+\mu}{\sigma^2}+2,\frac{z}{z-1}\right) \\
&- 2\left(\mu^2-\mu q+2\rho(\sigma^2+q)\right)\log(z)\Big)\Big/\left(2\rho(\mu^2-\mu q+2\rho(\sigma^2+q))\right),
\end{aligned}
$$

$q=\sqrt{\mu^2+2\rho\sigma^2}$ and ${}_2F_1$ is the hypergeometric function.

It is easy to check that $w(0)=\mu/\rho^2$, corresponding to the classical Lucas case with log utility and a risk-free sector in zero net supply. We also have $w(1)=0$, corresponding to the case when the whole economy is risk-free, and $w(0)=\int_0^\infty e^{-\rho s}\log(1)ds=0$.

For $\gamma>1$,

$$
\begin{aligned}
(1-\gamma)w(z) &= E\left[\int_0^\infty e^{-\rho s}\left(B_0\Big(1+\frac{D_0}{B_0}e^{y(s)}\Big)\right)^{1-\gamma}ds\right] \\
&= (1-z)^{1-\gamma}E\left[\int_0^\infty e^{-\rho s}\left(1+\frac{z}{1-z}e^{y(s)}\right)^{1-\gamma}ds\right] \\
&= \int_{-\infty}^\infty e^{\frac{y\left(\mu-\sqrt{\mu^2+2\rho\sigma^2}\,\mathrm{sign}(y)\right)}{\sigma^2}}\frac{\left(1+\frac{z}{1-z}e^y\right)^{1-\gamma}}{q} \\
&= \frac{(1-z)^{1-\gamma}}{q} \\
&\times \Bigg[\left(\frac{1-z}{z}\right)^{\frac{\mu-q}{\sigma^2}}\left(V\left(\frac{1-z}{z},\gamma+\frac{q-\mu}{\sigma^2},1-\gamma\right)+V\left(\frac{1-z}{z},\gamma+\frac{q-\mu}{\sigma^2}-1,1-\gamma\right)\right) \\
&+ \left(\frac{z}{1-z}\right)^{-\frac{q+\mu}{\sigma^2}}\left(V\left(\frac{z}{1-z},\frac{q+\mu}{\sigma^2},1-\gamma\right)+V\left(\frac{z}{1-z},\frac{q+\mu}{\sigma^2}+1,1-\gamma\right)\right)\Bigg],
\end{aligned}
$$

where $V(x,a,b)\overset{\text{def}}{=}\int_0^x t^{a-1}(1+t)^{b-1}dt$, and the result follows. ■

Proof of Proposition 3:

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

Now, from Equation (4), it follows that

$$
\begin{aligned}
\sqrt{2\pi\sigma^2\tau}\frac{P^\tau}{(1+e^d)^\gamma e^{-\rho\tau}} &= \int_{-\infty}^{\infty} \frac{e^{-(y-d-\mu\tau)^2/(2\sigma^2\tau)}}{(1+e^y)^\gamma} dy \\
&= \left(\int_{-\infty}^{-\epsilon} + \int_{\epsilon}^{\infty} + \int_{-\epsilon}^{\epsilon}\right) \frac{e^{-(y-d-\mu\tau)^2/(2\sigma^2\tau)}}{(1+e^y)^\gamma} dy \\
&= \int_{-\infty}^{-\epsilon} \frac{e^{-(y-d-\mu\tau)^2/(2\sigma^2\tau)}}{(1+e^y)^\gamma} dy + \int_{\epsilon}^{\infty} \frac{e^{-(y-d-\mu\tau)^2/(2\sigma^2\tau)}}{(1+e^y)^\gamma} dy + O(\epsilon) \\
&= \int_{-\infty}^{0} \frac{e^{-(y-\epsilon-d-\mu\tau)^2/(2\sigma^2\tau)}}{(1+e^{y-\epsilon})^\gamma} dy + \int_{0}^{\infty} \frac{e^{-(y+\epsilon-d-\mu\tau)^2/(2\sigma^2\tau)}}{(1+e^{y+\epsilon})^\gamma} dy + O(\epsilon)
\end{aligned} \tag{18}
$$

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

$$
\int_{-\infty}^{0} \sum_{n=0}^{\infty}(-1)^n a_n e^{-\frac{\epsilon}{2}n} e^{-(y-\epsilon-d-\mu\tau)^2/(2\sigma^2\tau)+n(y-\frac{\epsilon}{2})} dy. \tag{19}
$$

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 (18), 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 (11).

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 (19) to get

$$
\begin{aligned}
P^\tau &= \frac{(1+e^d)^\gamma e^{-\rho\tau-(d+\mu\tau)^2/(2\sigma^2\tau)}}{2}\times\\
&\sum_{n=0}^{\infty}(-1)^n a_n\left(F\left(\frac{d+\mu\tau+n\tau\sigma^2}{\sqrt{2\sigma^2\tau}}\right)+F\left(\frac{-d-\mu\tau+(n+\gamma)\tau\sigma^2}{\sqrt{2\sigma^2\tau}}\right)\right), \qquad (20)
\end{aligned}
$$

(iii): For

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

Equation (20) 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

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

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 4:
The result for r_s is standard. From Proposition 1 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:

$$\begin{aligned}\frac{1}{\sqrt{2\pi}}&\int_{-\infty}^{\infty}\frac{e^{-x^2/2}}{(1+e^d e^{x\sigma\sqrt{\tau}+\mu\tau})^\gamma}dx \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-\frac{\mu\tau+d}{\sigma\sqrt{\tau}}}\frac{e^{-x^2/2}}{(1+e^{x\sigma\sqrt{\tau}+\mu\tau+d})^\gamma}dx + \frac{1}{\sqrt{2\pi}}\int_{-\frac{\mu\tau+d}{\sigma\sqrt{\tau}}}^{\infty}\frac{e^{-x^2/2}}{(1+e^{x\sigma\sqrt{\tau}+\mu\tau+d})^\gamma}dx. \qquad (21)\end{aligned}$$

We prove the results for r^l by studying the first and second term in Equation (21) 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

$$N(-v) = C_2\frac{e^{-v^2/2}}{v}, \qquad C_2 \in \frac{1}{\sqrt{2\pi}}\left[\frac{v^2}{1+v^2}, 1\right], \tag{22}$$

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 (21), 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

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

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 (22).

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

$$\begin{aligned}
P^{\tau} &= (1+e^d)^{\gamma}e^{-\rho t}\left(\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-\frac{\mu\tau+d}{\sigma\sqrt{\tau}}}\frac{e^{-x^2/2}}{(1+e^{x\sigma\sqrt{\tau}+\mu\tau+d})^{\gamma}}dx+\frac{1}{\sqrt{2\pi}}\int_{-\frac{\mu\tau+d}{\sigma\sqrt{\tau}}}^{\infty}\frac{e^{-x^2/2}}{(1+e^{x\sigma\sqrt{\tau}+\mu\tau+d})^{\gamma}}dx\right)\\
&= (1+e^d)^{\gamma}e^{-\rho t}\left(C_3\frac{e^{-\frac{\mu^2}{2\sigma^2}\tau-\frac{\mu d}{\sigma^2}-\frac{d^2}{2\sigma^2\tau}}}{q}+C_4e^{-\frac{d^2}{2\sigma^2\tau}-\frac{d\mu}{\sigma^2}}\times\frac{e^{\frac{-\mu^2}{2\sigma^2}\tau}}{q_2}\right)\\
&= e^{-\left(\rho+\frac{\mu^2}{2\sigma^2}\right)\tau}(1+e^d)^{\gamma}e^{-\frac{\mu d}{\sigma^2}-\frac{d^2}{2\sigma^2\tau}}\left(\frac{C_3}{q}+\frac{C_4}{q_2}\right).
\end{aligned}$$

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 $\int_{-\infty}^{-\frac{\mu\tau+d}{\sigma\sqrt{\tau}}}\frac{e^{-x^2/2}}{(1+e^de^{x\sigma\sqrt{\tau}+\mu\tau})^{\gamma}}dx$-term in Equation (21) 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 $c_1e^{-\alpha_1\tau}+c_2e^{-\alpha_2\tau}\sim e^{-\min(\alpha_1,\alpha_2)\tau}$ 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

$$
-\frac{\log(P^{\tau})}{\tau} = -\frac{1}{\tau}\log\left((1+e^d)^{\gamma}e^{-\rho\tau}\left(e^{-\tau\frac{\mu^2}{2\sigma^2}}e^{-\frac{\mu d}{\sigma^2}-\frac{d^2}{2\sigma^2\tau}}\frac{C_3}{q} + C_5 e^{-\tau(\gamma\mu-\gamma^2\sigma^2/2)}\right)\right).
$$

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

$$
-\frac{\log(P^{\tau})}{\tau} = -\frac{1}{\tau}\left(\log\left((1+e^d)^{\gamma}e^{-\rho\tau}C_5 e^{-\tau(\gamma\mu-\gamma^2\sigma^2/2)}\right)+o(\tau)\right) = \frac{(\rho+\gamma\mu-\gamma^2\sigma^2/2)\tau+o(\tau)}{\tau},
$$

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 5: When $\mu < \gamma\sigma^2$, we have $v(z) \stackrel{\text{def}}{=} r^l - r^s = \frac{\mu^2}{2\sigma^2} - (\mu+\sigma^2/2)\gamma z + \gamma(\gamma+1)\frac{\sigma^2}{2}z^2$.

We have, $v''(z) = \gamma(\gamma+1)\frac{\sigma^2}{>}0$, so this is a convex function. Moreover, $v'(z) = 0$ at $z = \frac{\mu}{\sigma^2}+\frac{1}{2}$, so (iii) and (iv) follow immediately.

Now, the solutions to $v(z) = 0$ are $z = \frac{\mu}{\sigma^2}+\frac{1}{2}\pm q$. To ensure that the solutions are real valued, we define $C = \frac{\mu}{\sigma^2} < \gamma$. We then have $q^2 = (C+\frac{1}{2})^2 - \frac{1+\gamma}{\gamma}C^2 = \frac{1}{\gamma}\left(\frac{1}{4}+C(1-\gamma C)\right) > 0$, so indeed the solutions are real valued.

We also note in passing that $\mu < \gamma\sigma^2$ is also a sufficient condition for both the solutions to $v(z) = 0$ to lie strictly within the interval $[0,1]$.

∎

Proof of Corollary 1: Follows immediately from the arguments in Proposition 4.

∎

Proof of Corollary 2: Follows immediately from the arguments in Proposition 4. ■

Proof of Corollary 3: Follows immediately from the arguments in Proposition 4. ■

Mathematica code

We have verified numerically that the formulae for the long rate given in Proposition 4 are indeed correct, by directly evaluating Equation (4). 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 (13). 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 (14) 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

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

Ang, A., G. Bekaert, and M. Wei (2007): "The Term Structure of Real Rates and Expected Inflation," Working Paper 12930, NBER.

Ang, A., M. Piazzesi, and M. Wei (2006): "What does the Yield Curve tell us about GDP Growth?," *Journal of Econometrics*, 131, 359–403.

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

Bernard, H., and S. Gerlach (1998): "Does the Term Structure Predict Recessions? The International Evidence," *International Journal of Finance and Economics*, 3, 195–215.

Bonser-Neal, C., and T. R. Morley (1997): "Does the Yield Spread Predict Real Economic Activity? A Multicountry Analysis," *Federal Reserve Bank of Kansas City Economic Review*, pp. 37–53.

Brealey, R. A., S. C. Myers, and F. Allen (2008): *Principles of Corporate Finance.* McGraw-Hill, 9 edn.

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.

Chen, N. F. (1991): "Financial Investment Opportunities and the Macroeconomy," *Journal of Finance*, 46, 529–554.

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

Dotsey, M. (1998): "The Predictive Content of the Interest Rate Term Spread for Future Economic Growth," *Federal Reserve Bank of Richmond Economic Quarterly*, 84(3), 31–51.

Dybvig, P., J. Ingersoll, and R. Ross (1996): "Long forward and zero-coupon rates can never fall," *Journal of Business*, 69, 1–25.

Egorov, Y., and M. A. Shubin (1992): *Partial Differential Equations I.* Springer, Berlin.

Estrella, A., and G. A. Hardouvelis (1991): "The Term Structure as a Predictor of Real Economic Activity," *Journal of Finance*, 46, 555–576.

Estrella, A., and F. S. Mishkin (1997): "The Predictive Power of the Term Structure of Interest Rates in Europe and the United States: Implications for the European Central Bank," *European Economic Review*, 41, 1375–1401.

——— (1998): "Predicting U.S. Recessions: Financial Variables as Leading Indicators," *Review of Economics and Statistics*, 1, 45–61.

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.

Gustafsson, B., H. O. Kreiss, and J. Oliger (1995): *Time Dependent Problems and Difference Methods*. Wiley.

Hamilton, J. D., and D. H. Kim (2002): "A Re-examination of the Predictability of the Yield Spread for Real Economic Activity," *Journal of Money, Credit and Banking*, 34, 340–360.

Harvey, C. R. (1986): "Recovering Expectations of Consumption Growth from an Equilibrium Model of the Term Structure of Interest Rates," Ph.D. thesis, University of Chicago.

——— (1988): "The Real Term Structure and Consumption Growth," *Journal of Financial Economics*, 22, 305–334.

——— (1989): "Forecasts of Economic Growth from the Bond and Stock Markets," *Financial Analysts Journal*, 45, 38–45.

——— (1991): "The Term Structure and World Economic Growth," *Journal of Fixed Income*, 1, 7–9.

——— (1993): "The Term Structure Forecasts Economic Growth," *Financial Analysts Journal*, 49, 6–8.

Jorion, P., and F. S. Mishkin (1991): "A Multi-Country Comparison of Term Structure Forecasts at Long Horizons," *Journal of Financial Economics*, 29, 59–80.

Kessel, R. A. (1965): "The Cyclical Behavior of the Term Structure of Interest Rates," Occasional paper 91, NBER, New York, NY.

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.

Kreiss, H. O., and J. Lorenz (1989): *Initial-Boundary Value Problems and the Navier-Stokes Equations*. Academic Press, San Diego.

Laurent, R. (1988): "An Interest Rate-Based Indicator of Monetary Policy, Economic Perspectives," *Economic Perspectives*, 12, 3–14.

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

Majumdar, M., and N. Hashimzade (2006): "Survival, Uncertainty and Equilibrium Theory: An Exposition," Working paper, Cornell University, Department of Economics.

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

Moody, A., and M. P. Taylor (2003): "The High Yield Spread as a Predictor of Real Economic Activity: Evidence of a Financial Accelerator for the United States," Working paper, IMF.

Nelson, C. R., and A. F. Siegel (1987): "Parsimonious Modeling of Yield Curves," *Journal of Business*, 60, 473–489.

Parker, J. (2001): "The consumption risk of the stock market," *Brookings papers on economic activity*, 2, 279–348.

Parlour, C., R. Stanton, and J. Walden (2009): "Banking and Asset Prices," Working paper, UC Berkeley, Haas School of Business.

Plosser, C. I., and K. G. Rouwenhorst (1994): "International Term Structures and Real Economic Growth," *Journal of Monetary Economics*, 33, 133–155.

Rietz, T. (1988): "The equity risk premium," *Journal of Monetary Economics*, 22, 117–131.

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

Santos, T., and P. Veronesi (2006): "Labor Income and Predictable Stock Returns," *Review of Financial Studies*, 19, 1–44.

Stock, J. H., and M. W. Watson (1989): "New Indices of Coincident and Leading Indicators," in *1989 NBER Macroeconomic Annual*, ed. by J. B. Taylor, and M. Woodford. MIT Press, Cambridge, MA.

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.

——— (2005): “Risk, Uncertainty and Asset Pricing Puzzles,” Working paper, Harvard University, Department of Economics.