Title
Nonexponential Discounting: A Direct Test And Perhaps A New Puzzle

Authors
Startz, Richard
Tsang, Kwok Ping

Publication Date
2012-07-15

Nonexponential Discounting: A Direct Test And Perhaps A New Puzzle

Richard Startz
University of California, Santa Barbara

Kwok Ping Tsang[*]
Virginia Tech

July 2012

Abstract

Standard models of intertemporal utility maximization assume that agents discount future utility flows at a constant rate—exponential discounting. Euler equations estimated over different time horizons should have equal discount rates but they do not. Rising term yield premia imply discount rates that rise with longer horizons since uncertainty is much too small to account for the difference in interest rates. Such deviations from exponential discounting are large enough to make a significant difference in consumption choices over long horizons. Our results can be viewed as providing estimates of horizon-specific discounts, or as a further puzzle concerning intertemporal substitution and uncertainty.

Keywords: intertemporal consumer choice, discounting, hyperbolic discounting, consumption, portfolio puzzles, CAPM.

JEL Classifications: D11, D91, E21

Introduction

[*] We are grateful for comments from Shelly Lundberg and John Campbell, and for support from the Cecil and Jane Castor Endowment at the University of Washington. Thanks to the New York Federal Reserve for having the first author as a guest, though the views here do not represent the views of either the NY Fed or the Federal Reserve system. Further thanks go to seminar participants at the NY Fed, UC Santa Barbara, and Academia Sinica. Please address all correspondence to the first author at startz@econ.ucsb.edu. Authors' addresses: Dept. of Economics, 217 North Hall, University of California, Santa Barbara, CA 93106-9210 and Department of Economics, Virginia Tech, 3032 Pamplin Hall (0316), Blacksburg, VA 24061, byront@vt.edu.

In the canonical model of dynamic, intertemporal utility maximization, agents are assumed to discount future flow utility exponentially so that the contribution of consumption τ periods in the future to today's utility is $e^{-\delta(\tau-t)}EU(C_{t+\tau})$. If the future is not discounted exponentially, much of what we think we know about intertemporal dynamics in macroeconomics and in finance is open to question. In his seminal article, Strotz shows that this assumption is necessary for dynamic consistency: "An individual who … does not discount all future pleasures at a constant rate … finds himself continuously repudiating his past plans" (1955, 173). Strotz's argument is so strong that almost all work involving intertemporal choice assumes exponential discounting, the well-known exception being the literature on what is generically called hyperbolic discounting.

We estimate the standard model of the consumption/investment tradeoff, but nonstandardly estimate the tradeoff across various time horizons, allowing us to ask directly whether or not future pleasures are discounted at a constant rate.[1] We ask whether agents exponentially discount utility across significant horizons, in particular comparing one-quarter decisions to five-year decisions. If utility one quarter away is discounted by $e^{-.25\times\delta^{(.25)}}$ and utility five years out is discounted by $e^{-5\times\delta^{(5)}}$, does $\delta^{(.25)}=\delta^{(5)}$? The answer appears to be no. What is more, at long horizons the estimated difference between $\delta^{(.25)}$ and $\delta^{(5)}$ is large enough to be economically important. Our approach is completely standard except that we look at tradeoffs over long horizons directly instead of cumulating sequences of one-period decisions over several years. In principle, such tradeoffs depend on both long-horizon risk and expected return. To give away the punch line of the paper, empirically risk turns out to be basically irrelevant, and the higher

[1] Parker and Julliard (2005) also look at consumption and Euler equation at long horizons, though their focus is the consumption-based CAPM.

expected returns available on average for longer holding periods imply higher long-horizon discount rates.

Preferences in intertemporal tradeoffs across various horizons are revealed by agents' decisions to invest in (relatively safe) government bonds of corresponding maturities. Yields are typically higher on longer-maturity bonds. These increasing yields, the yield premia, imply increasing discount rates, the discount premia. In other words, $\delta^{(5)} > \delta^{(.25)}$. This may seem to be a surprising claim, seeming to turn on its head the customary procedure in which one maintains the assumption of exponential discounting and attributes higher long-term yields to greater risk. Perhaps it should not be surprising, since at least since Backus, Gregory, and Zin (1989) it has been known that the customary procedure has great difficulty in accounting for term premia without abandoning the additively separable, expected present value of utility model. As a matter of form, our statistical estimates allow for both nonexponential discounting and a response to risk, but as a practical matter the effect of the former overwhelms the latter. When long-term government bonds are held to maturity, the only relevant sources of uncertainty are unanticipated inflation and surprises that affect desired consumption growth. Conditional on these two sources of uncertainty, the risk of short-run fluctuations in long-run bond prices does not directly enter the consumer's optimization.[2] Neither source of uncertainty is very large, and greater uncertainty over long horizons rather than over short horizons increases estimates of the deviation from exponential discounting. Risk preferences do matter in these calculations, though less than might be expected because that part of uncertainty that evolves according to a random walk drops out

[2] Long bonds are usually considered risky because the one-period holding period return can be volatile. Nothing in our model is inconsistent with looking at one-period returns, but in the canonical model a Euler equation of any horizon is equally valid.

of the model. We use Weil's (1989) insight that positive discount rates constrain the coefficient of relative risk aversion to bound risk preferences.

We can summarize our results in two parts. The first finding is that the null of exponential discounting is soundly rejected. The second finding consists of estimates of an alternative, nonexponential discount schedule. We find that the distant future is discounted more heavily than is the near future.[3] The difference is large enough that projections of short-run tradeoffs to long-run decisions are likely to lead to considerable error. In the next section we work through this intuition quantitatively, presenting a calibration based on long-run averages. The following section provides econometric estimates based on the usual Euler equation estimation. We then turn to several robustness checks.

We estimate the standard model of intertemporal substitution for a representative consumer using aggregate data. Usually these models look at one-period growth in marginal utility set against one-period returns, r_t, and the discount rate δ, $E\left[\frac{[e^{(r_t-\delta)}U]'(C_{t+1})}{U'(C_t)}\right] = \mathbf{1}$. Our departure from the norm is simply to look at horizon t versus $t+m$ as well as at t versus $t+\mathbf{1}$.[4] Our estimates share all the usual merits and demerits of extracting preference parameters from aggregate data. One issue is significant: our estimates of the discount schedule are valid under the null of exponential discounting, but are not necessarily valid under the alternative. The

[3] We find that discount rates between periods at distant horizons are higher than discount rates between periods at near horizons, a result at odds with the experimental literature on nonexponential discounting (see Angeletos et al., 2001, for a long list of references). For example, Laibson (1997) writes, "Research on animal and human behavior has led psychologists to conclude that discount functions are approximately hyperbolic (Ainslie 1992). Hyperbolic discount functions are characterized by a relatively high discount rate over short horizons and a relatively low discount rate over long horizons" (445). Angeletos et al. (2001, 50) reiterate Laibson's (1997) statement: "The experimental evidence implies that the actual discount function declines at a greater rate in the short run than in the long run." The difference between our discount estimates and quasi-hyperbolic discounting is discussed below.

[4] In his *Handbook* chapter Singleton (1990) estimates the same equation that we do, but for very short horizons. He is not looking for nonexponential discounting, and the discount rates he reports in his Table 12.6 do not appear to be significantly different from one another. Note, however, that the large changes he finds in the estimates of the intertemporal elasticity of substitution may be picking up a nonexponential discounting effect.

distinction has more bite here than in many situations. We estimate the standard Euler equation across different horizons, but if there is nonexponential discounting it is not at all clear that the standard Euler equations apply. When discounting is not exponential, the correct Euler equation can take the form derived from the quasi-geometric approach (see Laibson, 1997, or Krusell and Smith, 2003) as a game across different selves or, more recently, that derived from a preference-for-temptation framework in the style of Gul-Pesendorfer (2001, 2004, 2005).[5] A logically consistent reading of our results thus allows for discarding exponential discounting without necessarily accepting our alternative estimates of discount rates. Readers should refer to the aforementioned studies for discount schedules estimated from some correctly specified Euler equations.

In the body of the paper we report alternative estimates of discount rates and take them to be meaningful. Other interpretations are possible, and the evidence is more clear in rejecting the null than it is decisive in choosing among alternatives. One (distressing) interpretation is that the dynamic programming/Euler equation model, which is ubiquitous in short-run models in macroeconomics and finance, is a useful approximation pf preferences at short horizons but not at intermediate and longer horizons. Preferences that allow for dynamic programming are, after all, a subset of all possible preferences—albeit a very useful subset. Under the slightly more general Strotz (1956) and Pollak (1968) assumptions that the current period agent “honors” the preferences of later period agents, Barro (1999) has shown that under log utility, nonexponential discount rates should not show up in the data.[6] Since deviations from exponential discounting do

[5] See Krusell, Kuruşçu, and Smith (2008) for an application of that framework to the optimal taxation problem.
[6] Strotz (1956) and Pollak (1968) assume that in the current period the consumer solves a dynamic programming problem that accounts for the solution she will later solve with changed preferences. Barro (1999) writes, “With no commitment ability and log utility, the equilibrium exhibits a constant effective rate of time preference and is observationally equivalent to the standard model.” Luttmer and Mariotti (2003) show this result also applies “when endowments are such that expected utility growth is constant,” but not more generally. All this means that there are

show up in the data, one can ask if this is evidence against the Strotz/Pollak formulation, or perhaps time preferences are better represented by some version of the “preferred habitat” model of Modigliani and Sutch (1966, 1967).

A different approach is that we are seeing a failure to sort out time preference versus risk preference rather than short- versus long-maturity time preference. This would suggest that the alternative relates to the still-sought-after resolution to the equity premium puzzle (or perhaps better, the “risk-free rate puzzle”) in the sense that we rely on the canonical model and the equity premium puzzle raises doubts about that model (see Mehra and Prescott, 1985, 2003 or the survey by Kocherlakota, 1996). However, the measure of risk in the equity premium puzzle is the risk to one-period returns. We also measure the risk to longer-term tradeoffs and show that, perhaps unsurprisingly, long-term nominal bonds have relatively little long-term risk.[7] This suggests that an alternative needs to not only separate attitudes toward risk and return but also toward short-horizon and long-horizon risk. Having said that, we reestimate what we take to be the leading empirical contender in this area, the Piazzesi and Schneider (2007) implementation of a simplified version of Epstein-Zin preference, and are unable to account for the apparent maturity specific rates of time preference.

Obviously, this paper does not explore all possibilities, and our conclusion may be weakened due to their absence. Nonseparability between durables and nondurables, between

---

circumstances under which consumers have nonexponential discount rates that our method fails to reveal. Empirically, we do find nonexponential discounting, implying either that the Strotz/Pollak assumption does not apply or that utility is far enough from logarithmic to be detectable.

[7] Uncertainty also plays a role in the extent to which asset prices can reveal information about nonexponential preferences. If future returns are certain, then if the strong form of the expectations hypothesis does not hold the consumer will attempt to arbitrage between short and long rates without regard to her rate(s) of discount (Kocherlakota (2001). Slightly more generally, if either sequential short-term or long-term investments first-order stochastically dominate, then again the consumer will choose the superior return without regard to discount rates. Under these circumstances the yield curve provides no information about discount rates. Apparently, neither condition pertains in the data.

consumption and leisure, and many other issues can all lead to the rejection of the exponential framework. One notable case is allowing consumption growth to have a persistent mean and time-varying volatility. Alvarez and Jermann (2005) show that in order to account for the term premium (and other asset pricing facts), one needs a utility function that can magnify the importance of the small permanent component in consumption. In fact, the long-run risk approach proposed by Bansal and Yaron (2004), in which consumption growth has persistent level and volatility and the consumer has Epstein-Zin preference, is shown to be able to account for some important asset pricing phenomena. A recent work by Doh (2011) shows that the Bansal-Yaron framework, when combined with inflation-specific volatilities, can indeed fit the U.S. term structure well.

With this preamble, we present our estimates of nonexponential discount rates, recognizing that some readers may prefer to identify the results as a nonexponential discount rate puzzle.

Model and Economic Estimates

In the canonical model, at time t the representative agent maximizes discounted utility. Let C_{t+m} be real consumption m periods hence. The applicable discount rate is $\delta^{(m)}$, where we use the parenthetical (m) to distinguish a symbol applicable over an *m*-year horizon. Under exponential discounting, $\delta^{(m)} = \delta\ \forall m$.

The objective function is

$$\int_{m=0}^{m=\infty} e^{-\delta^{(m)}m} EU(C_{t+m})dm \qquad (1~~1~~)$$

The consumer can save one nominal dollar today, invest it at the certain m -period nominal interest rate today, $r_t^{(m)}$, and increase nominal spending m periods hence by $e^{r_t^{(m)}m}$. While the Euler solution characterizing optimal behavior is usually written for the one-period consumption tradeoff, it is equally valid for all horizons.[8] Hence,

$$E\left[\frac{U'(C_{t+m})\big/P_{t+m}}{U'(C_t)\big/P_t}\right] = e^{\left(\delta^{(m)} - r_t^{(m)}\right)m} \qquad (2\cancel{2})$$

Suppose we now assume CRRA felicity, $U(C) = \dfrac{C^{1-\alpha} - 1}{1-\alpha}$, $U'(C) = C^{-\alpha}$. If we let $g^{(m)}$ be consumption growth over m periods ($g_t^{(m)}m \equiv \log(C_{t+m}) - \log(C_t)$) and $\pi_t^{(m)}$ be inflation, then we can write equation (2)~~(2)~~ in the more specific form

$$E\left[\left(e^{g_t^{(m)}m}\right)^{-\alpha} \times e^{-\pi_t^{(m)}m}\right] = e^{\left(\delta^{(m)} - r_t^{(m)}\right)m} \qquad (3)$$

If the bracketed term in (3) is approximately lognormal, then the left-hand side is $e^{\mu+\frac{\sigma^2}{2}}$, where $\mu^{(m)}$ and $\sigma^2_{(m)}$ are the mean and variance of the process $-\alpha g_t^{(m)}m - \pi_t^{(m)}m$. Taking expectations in equation (3) and then taking logs, the relation between discount rates and interest rates is

$$\delta^{(m)} = -\alpha \mathbf{E}\left(g_t^{(m)}\right) + \left[r_t^{(m)} - \mathbf{E}\left(\pi_t^{(m)}\right)\right] + \frac{\sigma^2_{(m)}}{2m} \qquad (4\cancel{4})$$

Equation (4)~~(4)~~ enters our calculations in two ways, levels and differences. First, when m is one period ($m = 0.25$ in our quarterly data), equation (4)~~(4)~~ gives the relation between the discount rate δ , the real interest rate $r - \mathbf{E}(\pi)$, consumption growth g , and uncertainty σ^2.

[8] Campbell (1986) and Harvey (1988) both make use of this result. See also Singleton (1990).

Imposing the nonnegativity on the level of the left-hand side, $\delta > \mathbf{0}$, places a limit on the admissible values of α that are admissible in the equation. Second, we use the long-horizon versus short-horizon difference in equation (4)~~(4)~~ to compute the discount premium $\delta^{(m_L)} - \delta^{(m_S)}$, as in equation (5)~~(5)~~.

$$\left(\delta^{(m_L)} - \delta^{(m_S)}\right) = -\alpha\left(E\left(g_t^{(m_L)}\right) - E\left(g_t^{(m_S)}\right)\right) + \left(\left[r_t^{(m_L)} - \mathbf{E}\left(\pi_t^{(m_L)}\right)\right] - \left[r_t^{(m_S)} - \mathbf{E}\left(\pi_t^{(m_S)}\right)\right]\right) + \left(\frac{\sigma^2_{(m_L)}}{2m_L} - \frac{\sigma^2_{(m_S)}}{2m_S}\right) \quad (5\cancel{5})$$

Temporarily pretend there is no uncertainty and replace the expectations in equation (4)~~(4)~~ with their long-run averages, $\bar{g}$, $\bar{r}^{(m)}$, and $\bar{\pi}$. Setting $\delta^{(.25)} > \mathbf{0}$ and solving gives $\alpha < \frac{\bar{r}^{(.25)} - \bar{\pi}}{\bar{g}}$. Using the data in **Error! Reference source not found.** we find $\alpha < .807$.[9] If we continue to ignore uncertainty and use long-run averages, we calculate the discount premium as

$$\delta^{(m_L)} - \delta^{(m_S)} = -\alpha\left(\overline{g^{(m_L)}} - \overline{g^{(m_S)}}\right) + \left[\left(\overline{r^{(m_L)}} - \overline{r^{(m_S)}}\right) - \left(\overline{\pi^{(m_L)}} - \overline{\pi^{(m_S)}}\right)\right]$$

Because calculation of long-run average growth and inflation is independent of the horizon, e.g., $\overline{g^{(m_L)}} = \overline{g^{(m_S)}}$, absent uncertainty the discount premium equals the average yield premium, $\delta^{(m_L)} - \delta^{(m_S)} = \overline{r^{(m_L)}} - \overline{r^{(m_S)}}$. Again referring to **Error! Reference source not found.**, we find $\delta^{(5)} - \delta^{(.25)} = 0.0108$. As a comparison, note this is the same order of magnitude as the short-term real rate, $\bar{r}^{(.25)} - \bar{\pi} = 0.0177$.

One might expect these calculations to be substantially changed by the inclusion of uncertainty. The convention, after all, is to assume exponential discounting and explain yield premia by risk. This is not the case. Using the values from **Error! Reference source not found.** we plot $\sigma^2_{(m)}(\alpha) = \alpha^2 var\left(g^{(m)}m\right) + var\left(\pi^{(m)}m\right) + 2\alpha cov\left(g^{(m)}m, \pi^{(m)}m\right)$ in **Error! Reference**

[9] Quoting Kocherlakota (1996, 50): "Note that the risk free rate puzzle comes from the equity premium puzzle: there is a risk free rate puzzle only if α is required to be larger than one so as to match up with the high equity premium."

source not found..[10] For $\alpha = 3$, the effect of risk is adding thirteen basis points to the five-year versus one-quarter yield premium. Our estimate of the departure from exponential discounting is $\delta^{(5)} - \delta^{(.25)} = 0.0121$.[11] Inclusion of uncertainty does very little to our calculation of the discount premium; its only effect is to increase the estimate. If this seems surprising, note that to the extent log consumption and inflation follow random walks, uncertainty has *no* effect on the discount premium calculation since $\sigma^2_{(m)}$ will be proportional to m , which will be exactly cancelled by the m in the denominator in equation (5)~~(5)~~.

While it may be that what we are seeing is an additional piece of evidence regarding the equity premium puzzle, these estimates of uncertainty suggest otherwise. **Error! Reference source not found.** shows that for decisions with an *m*-year horizon the risk of an *m*-year bond is negligible, and in fact that the long bond is safer than the short bond. Thus if one regards the apparent departure from exponential discounting as a puzzle, it looks to be an addition to the existing puzzle list.

Is the departure from exponential discounting economically important? The answer necessarily depends on the application, but one metric is to compare levels of consumption five years in the future that provide equal contributions to utility, first assuming a $\delta^{(.25)}$ discount rate and then assuming a $\delta^{(5)}$ discount rate.

In the next section, for $\alpha = 0.5$ we estimate $\delta^{(.25)} = 5.19 \times 10^{-3}$ and $\delta^{(5)} = 1.65 \times 10^{-2}$. The discounted value of utility five years out is $(1+\delta)^{-5} \times \dfrac{C_{t+5}^{1-\alpha} - 1}{1-\alpha}$. Call the utility level that

[10] We report unconditional uncertainty measures in **Error! Reference source not found.** where one really wants moments conditional on the consumer's information set. It appears that even unconditional uncertainty is small enough to not matter.

[11] Including uncertainty makes equation (4)~~(4)~~ a quadratic function in α . Solving for the largest admissible α for $m = 0.25$ turns out to give the same bound as before. Including uncertainty, α has to be smaller than .812 (or larger than 534.5, a possibility we ignore).

compensates consumption values with different discount rates $C_{t+5}^{(5)}$ and $C_{t+5}^{(.25)}$. Equation (7)~~(7)~~ gives the former in terms of the latter.

$$\left(C_{t+5}^{(5)}\right)^{1-\alpha} = \mathbf{1} + \left(\frac{1+\delta^{(.25)}}{1+\delta^{(5)}}\right)^{-5}\left(\left(C_{t+5}^{(.25)}\right)^{1-\alpha} - \mathbf{1}\right) \qquad (7\cancel{7})$$

If we solve equation (7)~~(7)~~ setting $C_{t+5}^{(.25)} = 13{,}628$ (mean consumption in our sample), we find $C_{t+5}^{(5)} = 15{,}227$, a 12 percent difference, which we consider moderately sized. Extrapolating by assuming a flat discount premium outside the range of our data, $\delta^{(30)} = \delta^{(5)}$, we would compute $C_{t+30}^{(5)} = 26{,}538$, just about twice the consumption predicted using exponential discounting based on the one-quarter discount rate. We consider this a large effect. The difference is smaller at shorter horizons both because the discount rate is small and because of the shorter compounding period. The compounding effect is greater at longer horizons.

In summary, we find that longer-horizon discount rates are relatively much higher than the short-horizon rate, but that the absolute level of the discount rate is low. The result is that discounting is of modest importance at short horizons and that the estimated deviation from exponential discounting is of considerable importance at longer horizons. In the next section we move from calibration to direct estimates of the Euler equations.

Estimation

The left-hand side of equation (2)~~(2)~~ plus a random error is observable. For a given value of α , the discount rate in the moment condition (8) can be estimated by least squares.[12] Since

[12] In principle, GMM can be used to estimate α and $\delta^{(m)}$ together, but one can get a wide variety of $\alpha's$ no apostrophe here or in next sentence, and change the "s" to Times to make it clear it's a plural. depending on the

our yields are not continuously compounded, equation (8) uses powers rather than the mathematically more-convenient exponential formulation.

$$E\left[\left(\frac{C_{t+m}}{C_t}\right)^{-\alpha} \times \frac{P_t}{P_{t+m}} \times \left(1+r_t^{(m)}\right)^m\right] = (1+\delta^{(m)})^m \tag{8}$$

Error! Reference source not found. shows estimates of the discount rate as a function of α for the usual one-period Euler equation.[13] Admissible ($\delta > \mathbf{0}$) values of α are somewhat lower than those found from long-run means in the previous section. Using the point estimates, $\alpha = 0.\mathbf{7}$ is the highest value for which $\delta > \mathbf{0}$. Values of $\alpha > 1.\mathbf{1}$ imply strictly negative confidence intervals for $\delta^{(.25)}$.

Error! Reference source not found. shows the (visually nearly indistinguishable) estimates of the discount premium, $\delta^{(m)} - \delta^{(.25)}$ for $\alpha = 0.5, 1,$ **and** $\mathbf{3}$. The premium rises to about 0.011 at five years, which is essentially the same number found in the introduction. Notably, estimates of the discount premium are unaffected by whether the discount rate itself is positive or negative. **Error! Reference source not found.** also provides confidence intervals, which are somewhat wider at longer maturities and notably wider for successively greater values of α. However, all the estimated discount premia are statistically significant.

The dashed line in **Error! Reference source not found.** shows the mean yield premia. As in the introduction, we find that discount premia estimates are essentially equal to the yield premia. Uncertainty in consumption and inflation is too small to make much difference.

---

instrument set. See Neely, Roy, and Whiteman (2001) and Yogo (2004) for the difficulty of identifying the EIS in the GMM context. Instead, we adopt the strategy of estimating $\delta^{(m)}$ for varying calibrated $\alpha's$. Because we want standard errors to account for serial correlation and to take advantage of contemporaneous correlation across horizons, we estimate using GMM with Newey-West standard errors. Point estimates are very close to least squares estimates.

[13] Confidence intervals are based on Newey-West standard errors throughout.

Formally, the null of exponential discounting is $\delta = \delta^{(m)} \forall m$. For values of α in the admissible range, exponential discounting is completely rejected. For $\alpha \leq 1$, the *p*-value is zero to all reported digits. At $\alpha = 2$ the *p*-value is 0.0024. Even at $\alpha = 3$ exponential discounting is rejected, though the rejection is only weakly significant ($p = 0.0875$).[14]

Phelps and Pollak (1968) and Laibson (1997) (see also Angeletos et al., 2001) suggest "quasi-hyperbolic" discounting as a specific departure from exponential. Quasi-hyperbolic discounts take the form $1, bd, bd^2, bd^3 \ldots$ with $d < 1$ and b substantially less than one. However, some of the important implications of quasi-hyperbolic discounting—Angeletos et al. (1991) use the term "salience of the present"—depend on b being substantially less than one. We fit our six estimated discount rates to $\left(1 + \widehat{\delta^{(m)}}\right)^{-m} = bd^m$ by nonlinear least squares. Given the picture in **Error! Reference source not found.** this gives an unsurprisingly near-perfect fit. We find $\hat{b} = 1.006$ with a standard error of 0.002. Since we do not find b to be substantially less than one, our estimates differ from quasi-hyperbolic specifications in an important way.[15]

Further considerations

In this section we look at several further considerations both of interpretation and for empirics.

Error! Reference source not found. provides an empirical estimate of discount premia. The associated tests reject the premia equaling zero. Our estimates are valid under the null of

[14] For *much* higher values of α the point estimates of discount premia are considerably larger, but are no longer statistically significant. For $\alpha = 30$, the estimate of the five-year premium is 0.0311.

[15] Rubinstein (2003) also raises doubts about quasi-hyperbolic discounting, albeit for quite different reasons.

exponential discounting, which is all the rejection requires. The discount premia estimates require that under the alternative hypothesis the Euler equations follow the usual variational argument at all horizons. However, once exponential discounting is abandoned, all sorts of alternative hypotheses can be and have been brought forward.

While we would not offer estimates of discount premia if we considered them uninteresting—the Euler equations applying at all horizons is clearly an interesting hypothesis—we cannot overemphasize the care needed in their interpretation. Dynamic inconsistency and other deviations in behavior from the canonical model are precisely the reasons that nonexponential discounting is so interesting. But, as an example, consider the finding in the previous section that b is too large to support quasi-hyperbolic discounting. There is a logically valid rejoinder that if choices are dynamically inconsistent, then Euler equations do not apply and the apparent deviations of the estimated discount schedule from the quasi-hyperbolic are not convincing.

This important point for interpretation having being made, we turn to further examination of the empirical estimates.

Longer Horizons

The longer the horizon, the greater the potential interest in nonexponential discounting. Using the CRSP zero coupon data set, we are limited to a five-year horizon.[16] Robert Bliss has

[16] While longer-term yields are available for coupon bonds, because of the coupons these yields do not exactly correspond to the yields that belong in an Euler equation. Additionally, we discard the last *m*-quarters of data in order to measure c_{t+m}, which is problematic for large *m*. With these caveats in mind, note that the yield premium over the five-year rate on constant maturity U.S. government bonds from 1977M2 through 2002M2 was twenty-four

computed an alternative set of zero coupon yields with a much richer set of maturities at both short and long horizons, albeit with shorter historical coverage.[17] Before using the longer Bliss maturities, we compare the common maturities in the Bliss and CRSP data sets for the Bliss period to show that the two are quite similar. Remembering that we estimated $\delta^{(.25)} = 5.19 \times 10^{-3}$ for $\alpha = 0.5$ for the full-sample CRSP data, our corresponding estimate using the CRSP data over the Bliss sample period is slightly higher at 6.64×10^{-3}. We find $\delta^{(5)} - \delta^{(.25)} = 1.53 \times 10^{-2}$ for the CRSP data, about 40 basis points higher than for the longer sample period. For the Bliss data we estimate $\delta^{(.25)} = 8.29 \times 10^{-3}$.

Error! Reference source not found. puts together CRSP and Bliss data estimates of the discount premia. The estimates are approximately equal for the one- through five-year horizons for which the datasets overlap. The Bliss data estimates continue to increase to the ten-year horizon, increasing with horizon to $\delta^{(10)} - \delta^{(.25)} = 2.59 \times 10^{-2}$. Using the utility compensating measure of consumption in equation (7)~~(7)~~ gives basically the same result as before for the five-year horizon (16 percent using CRSP data and 17 percent using Bliss data, as compared to 12 percent for the CRSP data period). The difference is much greater at longer horizons. Using the Bliss data, a 65 percent increase in consumption is required ten years out to give equivalent utility. Applying the ten-year discount rate at a thirty-year horizon, a quadrupling of consumption is required to match the utility flow using the one-quarter discount rate.

Confidence Intervals

basis points for the ten-year rate and thirty-nine basis points for the thirty-year rate (FRED II and authors' calculations).

[17] See Bliss (1997) and Fama and Bliss(1987). We are grateful to Robert Bliss for making his data available.

Confidence intervals in **Error! Reference source not found.** are based on Newey-West standard errors. With just under 200 observations, they are the equivalent of only ten nonoverlapping observations on five-year intervals, loosely speaking. As a check on our Newey-West inference, we compute two parametric bootstraps. First, we estimate the one-quarter Euler equation for $\alpha = 0.5$. We then generate artificial histories for $\frac{C_t^{-\alpha}}{P_t}$ using $\left(\frac{C_{t+.25}^{-\alpha}}{P_{t+.25}}\right) = \left(\frac{C_t^{-\alpha}}{P_t}\right)\left(1 + r_t^{(.25)}\right)^{-.25}\left[\left(1 + \widehat{\delta^{(.25)}}\right)^{.25} + e_t\right]$, where e_t is a residual from the estimated one-quarter Euler equation resampled with replacement. We then reestimate the system of parameters in **Error! Reference source not found.** one thousand times. This first bootstrap assumes the errors in the estimated one-quarter Euler equation are serially uncorrelated, as theory would suggest.[18] However, the empirical residuals are well modeled as an ARMA(1,1), $e_t = 0.88\mathbf{e_{t-1}} + \mathbf{\epsilon_t} - \mathbf{0.40\epsilon_{t-1}}$. As a second bootstrap we resample from the estimated innovations, $\mathbf{\epsilon_t}$, then generate both errors and artificial histories.

Error! Reference source not found. shows the original confidence region and both simulated $[0.025, 0.975]$ confidence intervals. The simulated intervals are notably tighter than the asymptotic intervals, reinforcing our confidence in the earlier statistical inference.

Habit Formation

As an extension to CRRA utility, which many writers have found attractive, consider Abel's (1999) external habit formation model. Abel replaces the period utility function with

[18] The Euler equation should be serially uncorrelated if measurement was perfect. Serial correlation may be induced due to both measurement error in the level of consumption and the Working capitalization necessary? effect. These would induce somewhat offsetting moving average errors. See Wilcox (1992) for a discussion.

$$U_t = \frac{1}{1-\alpha}\left(\frac{C_t}{v_t}\right)^{1-\alpha}$$
$$v_t \equiv C_t^{\boldsymbol{h_0}} C_{t-1}^{\boldsymbol{h_1}} \tag{9}$$

where v_t is an external reference standard. Equation (3) would be rewritten as

$$E\left[\left(e^{g_t^{(m)}m}\right)^{-\alpha} \times \left(e^{g_t^{(m)}m}\right)^{-(1-\alpha)\boldsymbol{h_0}} \times \left(e^{g_{t-1}^{(m)}m}\right)^{-(1-\alpha)\boldsymbol{h_1}} \times e^{-\pi^{(m)}m}\right] = e^{\left(\delta^{(m)}-r_t^{(m)}\right)m} \tag{10}$$

Taking logs, we see that habit formation adds a lagged variable that in principle could account for serial correlation in equation (3). Equation (4)~~(4)~~ becomes

$$\delta^{(m)} - r_t^{(m)} = -(\alpha + (1-\alpha)\boldsymbol{h_0})g_t^{(m)} - (1-\alpha)\boldsymbol{h_1} g_{t-1}^{(m)} - \pi^{(m)} + \frac{\sigma_{(m)}^2}{2m} \tag{11}$$

External habit formation may account for serial correlation and suggests that the behavioral parameter we identify as α might be $\alpha + (1-\alpha)(\boldsymbol{h_0} + \boldsymbol{h_1})$. However, for our long-run calculations $\overline{g_t^{(m)}} = \overline{g_{t-1}^{(m)}} = \overline{g^{(m)}}$, so external habit formation ought not affect the estimate of the discount premia. Since for our purposes identification of the behavioral parameters α , $\boldsymbol{h_0}$, and $\boldsymbol{h_1}$ is not needed, we modify equation (8)~~(8)~~ to include lagged consumption growth as in $E\left[\left(\frac{C_{t+m}}{C_t}\right)^{-a_1} \times \left(\frac{C_{t+m-1}}{C_{t-1}}\right)^{-a_2} \times \frac{P_t}{P_{t+m}} \times \left(1+r_t^{(m)}\right)^m\right] = (1+\delta^{(m)})^m$. Continuing with the strategy of fixing the left-hand side parameters, we estimate this version of the moment condition for $0 \le a_1 \le 3$ and $-3 \le a_2 \le 3$. Interestingly, for some values of a_2 habit formation improves the results in the sense of eliminating serial correlation and also giving higher estimates of $\delta^{(.25)}$ ("improves" in the sense that the results are closer to our usual priors).

Error! Reference source not found. shows habit formation results for $a_1 = 1.5$, which is the value most likely both to give $\delta^{(.25)} > 0$ and to eliminate serial correlation. The $a_2 = 0$ position on the horizontal axis corresponds to the no-habit-formation results given earlier. The top panel shows, as before, that the estimate of the one-quarter Euler equation gives a negative discount rate and the bottom panel shows there is significant serial correlation. In contrast, for habit formation with $a_2 < -1$ serial correlation disappears, and $\delta^{(25)}$ is positive and generally larger than was estimated without habit formation. Including habit formation improves the model, but the important conclusion for the purpose at hand is that habit formation makes no difference whatsoever in the estimate of the discount premium. It remains between 0.011 and 0.012, as can be seen in the middle panel.

Epstein-Zin Preference

Epstein and Zin (1991) offer a specific alternative to nonexponential discounting. In a widely cited application, Piazzesi and Schneider (2007) show that the Epstein-Zin recursive preference model is able to fit the slope of the yield curve with reasonable time preference and risk aversion parameters. The Epstein-Zin preference, with the CRRA preference as a special case, allows the reciprocal of the elasticity of intertemporal substitution (EIS) to be different from the relative risk aversion. Using a version of the Epstein-Zin preference in which the EIS is set to one, Piazzesi and Schneider show that the sample mean of two ends of the yield curve can be explained with discount factor β and risk aversion parameter γ of reasonable magnitude. Does the Epstein-Zin preference account for uncertainty in a way that eliminates the differential in maturity-specific discount rates?

To test for nonexponential discounting under the Epstein-Zin preference, we follow the estimation procedure of Piazzesi and Schneider with one modification (more details are available in the Appendix).[19] We first estimate a state-space model for log consumption growth and inflation, as in Piazzesi and Schneider (2007). Piazzesi and Schneider then fit the discount factor and risk aversion parameter by matching their model to the two ends of the yield curve. Since we want to allow for maturity-specific discount rates, we fix the risk aversion parameter γ (at the same value found in Piazzesi and Schneider and two other values) and choose the discount rate δ to fit the model mean of *each* yield to its sample mean. Since we get one discount rate estimate for each yield, we can see whether the discounting rates are maturity specific. As shown in Table 2, the estimated maturity-specific discount rates continue to have little difference from those we reported earlier, which tells us that the Epstein-Zin preference does not explain the nonexponential discounting problem.

Measurement Error

Measurement error in the level of consumption is a potentially important issue in Euler equation estimation (Altonji and Siow, 1987 and Singleton, 1990) since it can induce a moving average error in measured consumption that matters more for short-term than for long-term growth rates. From equation (5)~~(5)~~, one can see that in principle mismeasurement of uncertainty could matter. In practice, uncertainty is too small for mismeasurement to be an issue. As a check, we conduct a Monte Carlo experiment in which log consumption follows a random walk with mean growth as observed in the data, and a shock with one-half the growth variance in the data. We treat this series as simulated "true" consumption. We then add to the simulated *level* of

[19] We modify the programs provided on Piazzesi's website for our estimation. The original programs can be found at http://www.stanford.edu/~piazzesi/nberannualprograms.zip.

"true" consumption an error scaled to the other half of the variance, giving simulated "noisy" consumption.

We repeat our estimates on one thousand draws of the simulated data. **Error! Reference source not found.** shows median results for both "true" and "noisy" data. The two sets of simulation results are visually indistinguishable from one another—the difference is less than a ten of a basis point, which is smaller than the width of the lines in **Error! Reference source not found.**—and as a practical matter are indistinguishable from our actual results. This adds to our confidence that our discount premium estimates essentially reflect differences in yield premia with little effect of uncertainty.

Conclusion

Long-term, safe nominal yields are, on average, higher than short-term yields. In the canonical model of intertemporal expected utility maximization, a positive yield premium implies a positive discount premium and hence nonexponential discounting unless the difference is offset by risk. Empirically, uncertainty in consumption paths and inflation is small, and we find that discount rates rise with the decision horizon.

Exponential discounting is rejected statistically. The economic consequences of nonexponential discounting are modest over short horizons because discount rates are small. Over longer horizons, compounding makes the economic consequences much larger. Looking thirty years out, the difference in utility-equivalent consumption using a five-year rather than one-quarter discount rate is a factor of two; if one uses the estimate of a ten-year discount rate from the shorter Bliss data sample it is a factor of four.

Of course, all this is a joint test of exponential discounting and preference specification. Perhaps agents are not infinite-horizon, additive-expected utility maximizers, or perhaps aggregation does not give data in which the behavior of the average individual looks like the behavior of a representative agent.[20] If in addition to rejecting exponential discounting one is willing to accept our estimates of discount premia structure, then one has to accept the use of Euler equations despite the time-inconsistency issues. With that caveat in mind, the departure from exponential discounting is large enough to be critical to an understanding of the differences between short-horizon and long-horizon decision making.

References

Abel, Andrew B. (1999) "Risk Premia and Term Premia in General Equilibrium," *Journal of Monetary Economics* 43(1), 3-33.

Ainslie, George W. (1992) *Picoeconomics*, City: Cambridge University Press.

Altonji, Joseph, and Aloysius Siow. (1987) "Testing the Response of Consumption to Income Changes with (Noisy) Panel Data," *Quarterly Journal of Economics* 102(2), 293-328.

Alvarez, Fernando, and Urban J. Jermann. (2005) "Using Asset Prices to Measure the Persistence of the Marginal Utility of Wealth," *Econometrica* 73(6), 1977-2016.

Angeletos, George-Marios, David Laibson, Andrea Repetto, Jeremy Tobacman, and Stephen Weinberg. (2001) "The Hyperbolic Consumption Model: Calibration, Simulation, and Empirical Evaluation," *Journal of Economic Perspectives* 15(3), 47-68.

[20] A suggestion for future research from Shelly Lundberg is to consider whether rising hazard rates for mortality might induce nonexponential discounting without necessarily leading to time inconsistency.

Backus, David K., Allan W. Gregory, and Stanley E. Zin. (1989) "Risk Premiums in the Term Structure: Evidence from Artificial Economies," *Journal of Monetary Economics* 24(3), 371-99.

Bansal, Ravi, and Amir Yaron. (2004) "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance* 59(4), 1481-1509.

Barro, Robert. (1999) "Ramsey Meets Laibson in the Neoclassical Growth Model," *Quarterly Journal of Economics* 114(4), 1125-52.

Bliss, Robert R. (1997) "Testing Term Structure Estimation Methods," *Advances in Futures and Options Research*, vol. 9, City: Press.

Campbell, John Y. (1986) "Bond and Stock Returns in a Simple Exchange Model," *Quarterly Journal of Economics* 101(4), 785-804.

CRSP Data. (2008) Center for Research in Security Prices. Graduate School of Business, The University of Chicago.

Doh, Taeyoung. (2011) "Long-Run Risks in the Term Structure of Interest Rates: Estimation," Federal Reserve Bank of Kansas City, working paper RWP 08-11.

Epstein, Lawrence G., and Stanley E. Zin. (1991) "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis," *Journal of Political Economy* 99(2), 263−86.

Fama, Eugene, and Robert Bliss. (1987) "The Information in Long-Maturity Forward Rates," *American Economic Review*, 77(4), 680-92.

Gul, Faruk, and Wofgang Pesendorfer. (2001) "Temptation and Self-Control," *Econometrica* 69(6), 1403–35.

Gul, Faruk, and Wolfgang Pesendorfer. (2004) "Self-Control and the Theory of Consumption," *Econometrica* 72(1), 119–58.

Gul, Faruk, and Wolfgang Pesendorfer. (2005) "The Revealed Preference Theory of Changing Tastes," *Review of Economic Studies* 72(2), 429–48.

Gürkayank, Refet S., Brian Sack, and Jonathan H. Wright. (2008) "The TIPS Yield Curve and Inflation Compensation," Finance and Economics Discussion Series Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board, working paper, 2008-05.

Hall, Robert E. (1988) "Intertemporal Substitution in Consumption," *Journal of Political Economy* 96(2), 339-57.

Hansen, Lars P., and Kenneth J. Singleton. (1983) "Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns," *Journal of Political Economy* 91(2), 249-65.

Harvey, Campbell R. (1988) "The Real Term Structure and Consumption Growth," *Journal of Financial Economics* 22(2), 305-34.

Kocherlakota, Narayana R. (1996) "The Equity Premium: It's Still a Puzzle," *Journal of Economic Literature* 34(1), 42-71.

Kocherlakota, Narayana R. (2001) "Looking for Evidence of Time-Inconsistent Preferences in Asset Market Data," *Federal Reserve Bank of Minneapolis Quarterly Review* 25(3), 13–24.

Krusell, Per, and Anthony A. Smith, Jr. (2003) "Consumption-Savings Decisions with Quasi-Geometric Discounting," *Econometrica* 71(1), 365-75.

Krusell, Per, Burhanettin Kuruşçu, and Anthony A. Smith, Jr. (2010) "Temptation and Taxation," *Econometrica* 78(6), 2063-84.

Laibson, David. (1997) “Golden Eggs and Hyperbolic Discounting,” *Quarterly Journal of Economics* 112(2), 443–77.

Luttmer, Erzo G. J., and Thomas Mariotti. (2003) “Subjective Discounting in an Exchange Economy,” *Journal of Political Economy* 111(5), 959-89.

Mehra, Rajnis,h and Edward C. Prescott. (1985) “The Equity Premium: A Puzzle,” *Journal of Monetary Economics* 15(2), 145-61.

Mehra, Rajnish, and Edward C. Prescott. (2003) “The Equity Premium in Retrospect,” *Handbook of the Economics of Finance*, eds. G. M. Constantinides, M. Harris, and R. Stulz, City: Elsevier B.V.

Modigliani, Franco, and Richard Sutch. (1966) “Innovations in Interest Rate Policy,” *American Economic Review* 56(2), 178-97.

Modigliani, Franco, and Richard Sutch. (1967) “Debt Management and the Term Structure of Interest Rates: An Empirical Examination of Recent Experience,” *Journal of Political Economy* 75(4), 569-89.

Neely C. J., A. Roy, and C. H. Whiteman. (2001) “Risk Aversion Versus Intertemporal Substitution: A Case Study of Identification Failure in the Intertemporal Consumption Capital Asset Pricing Model,” *Journal of Business & Economic Statistics* 19(4), 1 395-403.

Parker, Jonathan A., and Christian Julliard. (2005) “Consumption Risk and the Cross Section of Expected Returns,” *Journal of Political Economy*,113(1), 185-222.

Phelps, Edmund S., and Robert A. Pollak. (1968) “On Second-Best National Saving and Game-Equilibrium Growth,” *Review of Economic Studies* 35(2), 185-99.

Piazzesi, Monika, and Martin Schneider. (2007) "Equilibrium Yield Curves," in Daron Acemoglu, Kenneth Rogoff, and Michael Woodford, eds., *NBER Macroeconomics Annual 2006*, Cambridge MA: MIT Press, 389-442.

Pollak, Robert A. (1968) "Consistent Planning," *Review of Economic Studies* 35(2), 201-208.

Rubinstein, Ariel. (2003) "Economics and Psychology? The Case of Hyperbolic Discounting," *International Economic Review* 44(4), 1207-16.

Singleton, Kenneth J. (1990) "Specification and Estimation of Intertemporal Asset Pricing Models," *Handbook of Monetary Economics*, vol. I, B.M. Friedman and F.H. Hahn, eds., City: Elsevier Science Publishers B.V.

Strotz, Robert H. (1955) "Myopia and Inconsistency in Dynamic Utility Maximization," *Review of Economic Studies* 23(3), 165-80.

Weil, Phillipe. (1989) "The Equity Premium Puzzle and the Risk-Free Rate Puzzle," *Journal of Monetary Economics* 24(3), 401-21.

Wilcox, David W. (1992) "The Construction of U.S. Consumption Data: Some Facts and Their Implications for Empirical Work," *American Economic Review* 82(4), 922-41.

Yogo, Motohiro. (2004) "Estimating the Elasticity of Intertemporal Substitution When Instruments Are Weak," *Review of Economics and Statistics* 86(3), 797-810.

Table 1: Sample Moments

Maturity in years (m)	$\bar{r}^{(m)}$	$var(g^{(m)}m)$	$var(\pi^{(m)}m)$	$cov(g^{(m)}, \pi^{(m)}m)$
0.25	5.36×10^{-2}	2.04×10^{-5}	5.29×10^{-5}	-1.04×10^{-5}
5	6.44×10^{-2}	9.92×10^{-4}	1.05×10^{-2}	-5.33×10^{-4}
	$\bar{g} = 0.0218$		$\bar{\pi} = 0.036$	

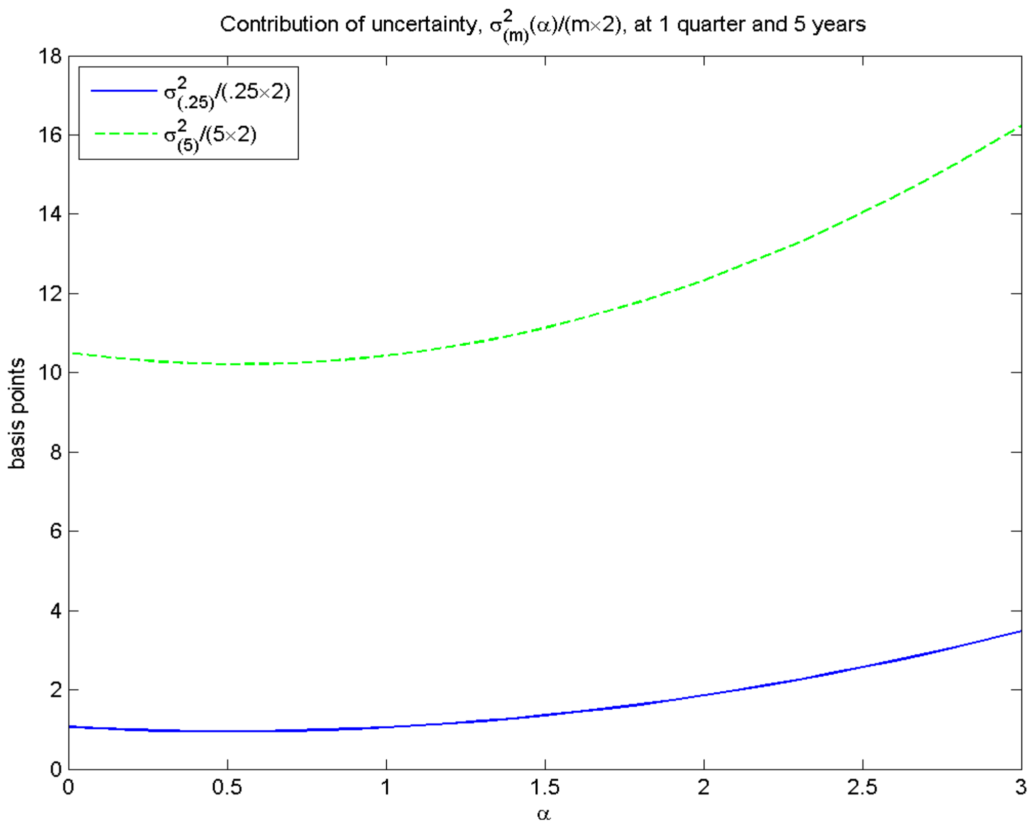
Note: Quarterly data for the period 1954 through 2002. The data are real per capita consumption of nondurables and services, a price index that is the weighted average of the consumption nondurables price index and the consumption services price index, the one-quarter treasury bill rate, and the one- through five-year CRSP zero coupon rates. Means are at annual rates., $g^{(m)}$ is the consumption growth over m periods ($g_t^{(m)}m \equiv \log(C_{t+m}) - \log(C_t)$), and $\pi_t^{(m)}$ is inflation over m periods.

Table 2: Calibration for Epstein-Zin Preference

	Risk Aversion Parameter		
Discount Rate (year)	$\gamma = 20$	$\gamma = 59$	$\gamma = 200$
$\delta^{(.25)}$	-0.0040	-0.0028	0.0013
$\delta^{(1)}$	0.0013	0.0025	0.0069
$\delta^{(2)}$	0.0041	0.0058	0.0124
$\delta^{(3)}$	0.0063	0.0087	0.0181
$\delta^{(4)}$	0.0082	0.0112	0.0228
$\delta^{(5)}$	0.0092	0.0128	0.0258

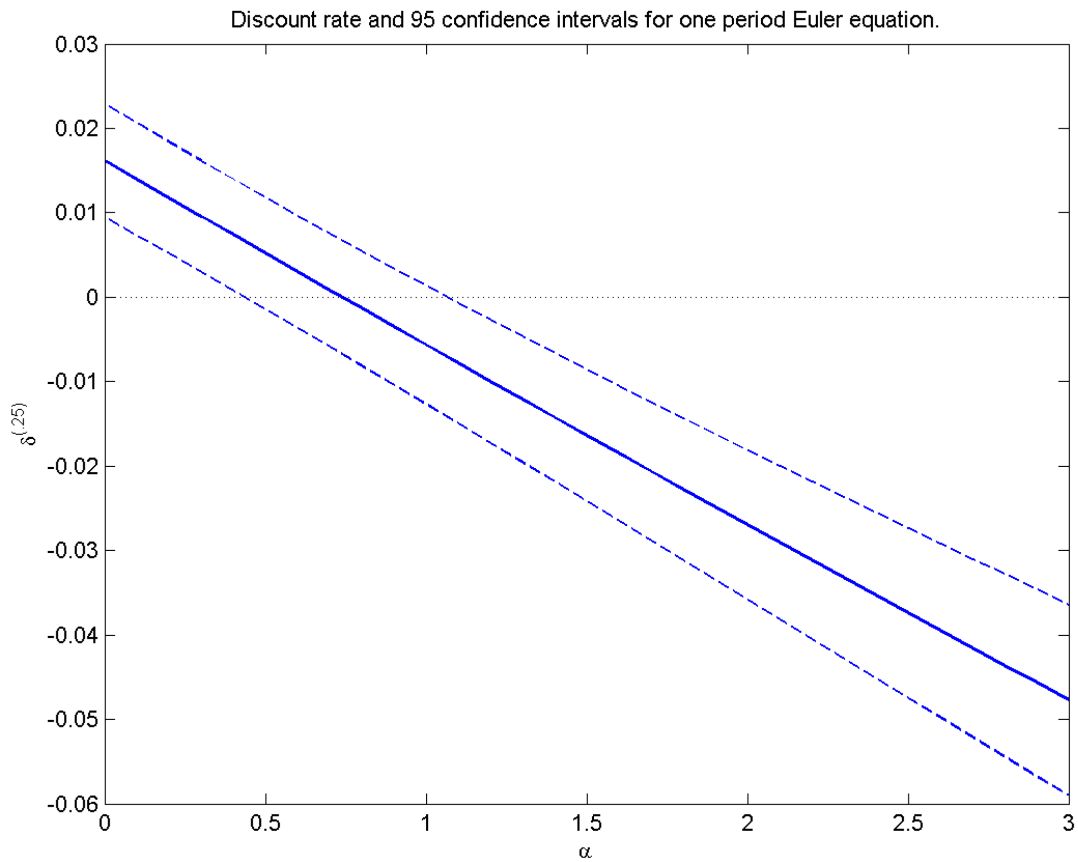
Note: Here we fix the risk aversion parameter γ (at the same value 59 used in Piazzesi and Schneider, 2007 and two other values) and choose the discount rate δ to fit the model mean of *each* yield to its sample mean. Please see section 4.4. for details.

Figure 1: Contribution of Uncertainty



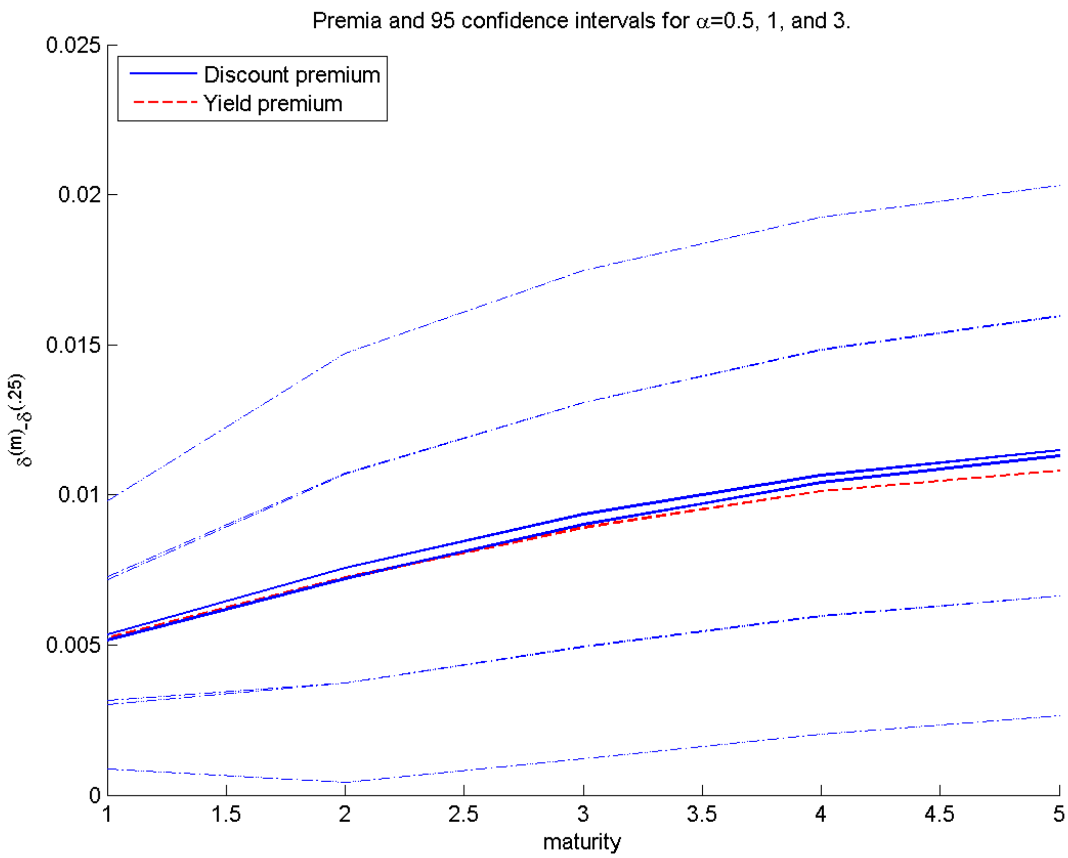
Note: We plot $\frac{\sigma^2_{(m)}(\alpha)}{2m}$ (in basis points) against different values of α, the parameter in the CRRA utility function. This corresponds to the last term in equation (4). For example, for $\alpha = 3$ the effect of risk is to add 13 basis points to the five-year versus one-quarter yield premium.

Figure 2:Discount Rate Implied by Different Values of α



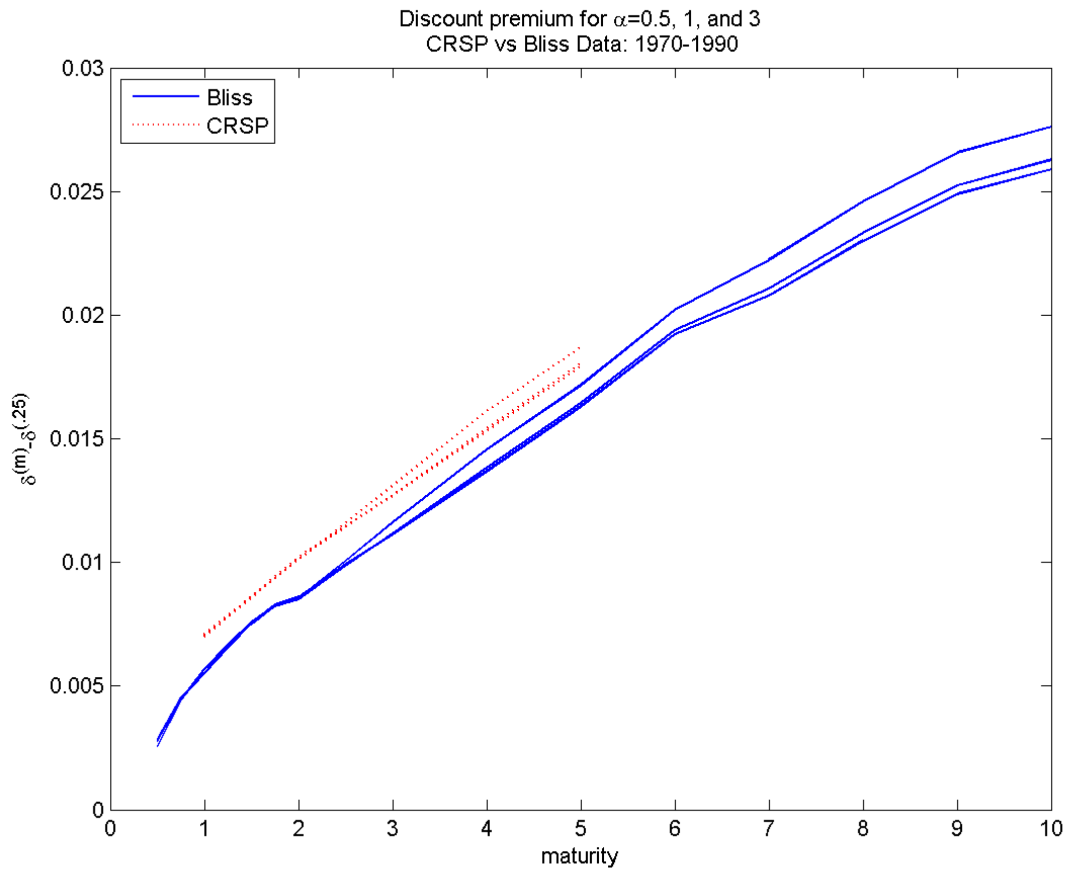
Note: We estimate the Euler equation by least squares by fixing the value of the CRRA parameter α . For each value of α , there is a different value for the implied discount rate δ . Here we show the estimates for using only one-quarter consumption growth and one-quarter bond yield.

Figure 3: Discount Premia For Different Maturities



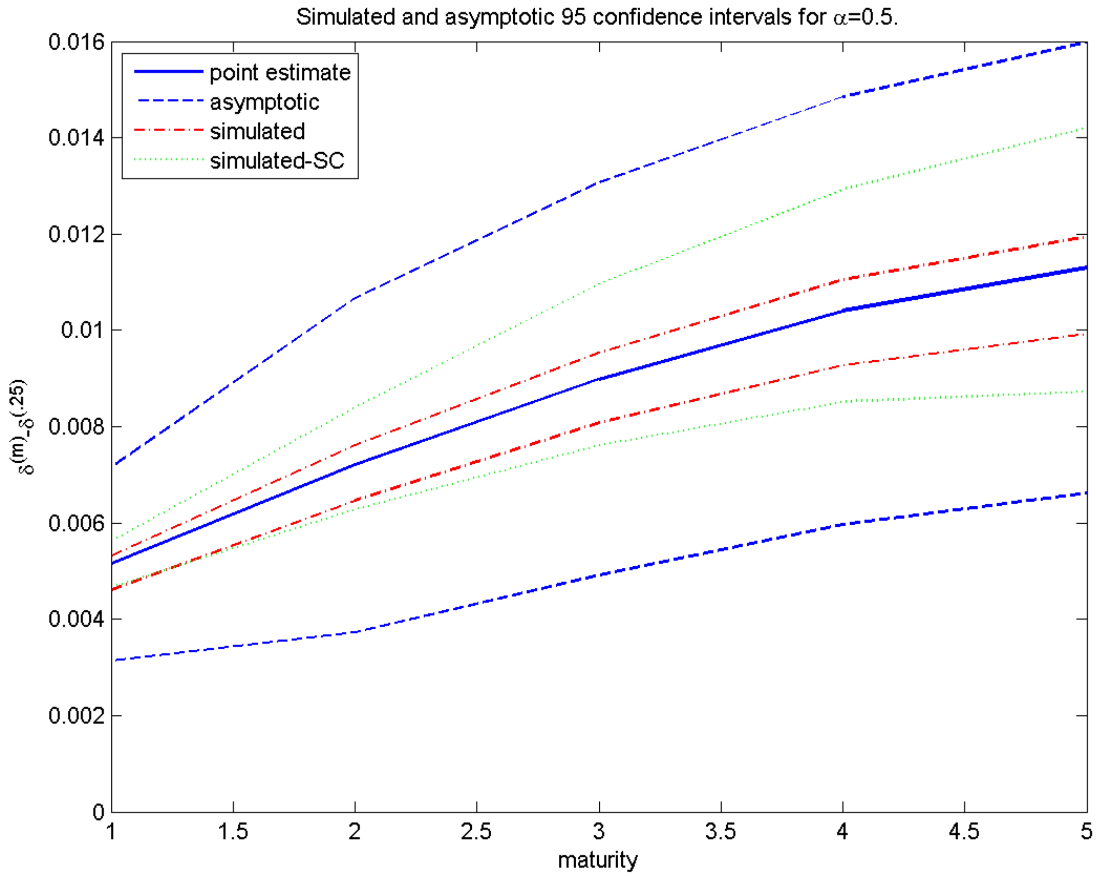
Note: Here we allow the discount rate to differ among maturities. We plot the difference between the m-year discount rate and the one-quarter discount rate against m. If the discount is exponential the blue line should be flat at zero. Since the estimates for the three values of α are very close (i.e., the three solid blue lines), the reader is referred to the Online Appendix for the exact point estimates.

Figure 4:Discount Premium for the Bliss (1997) Dataset



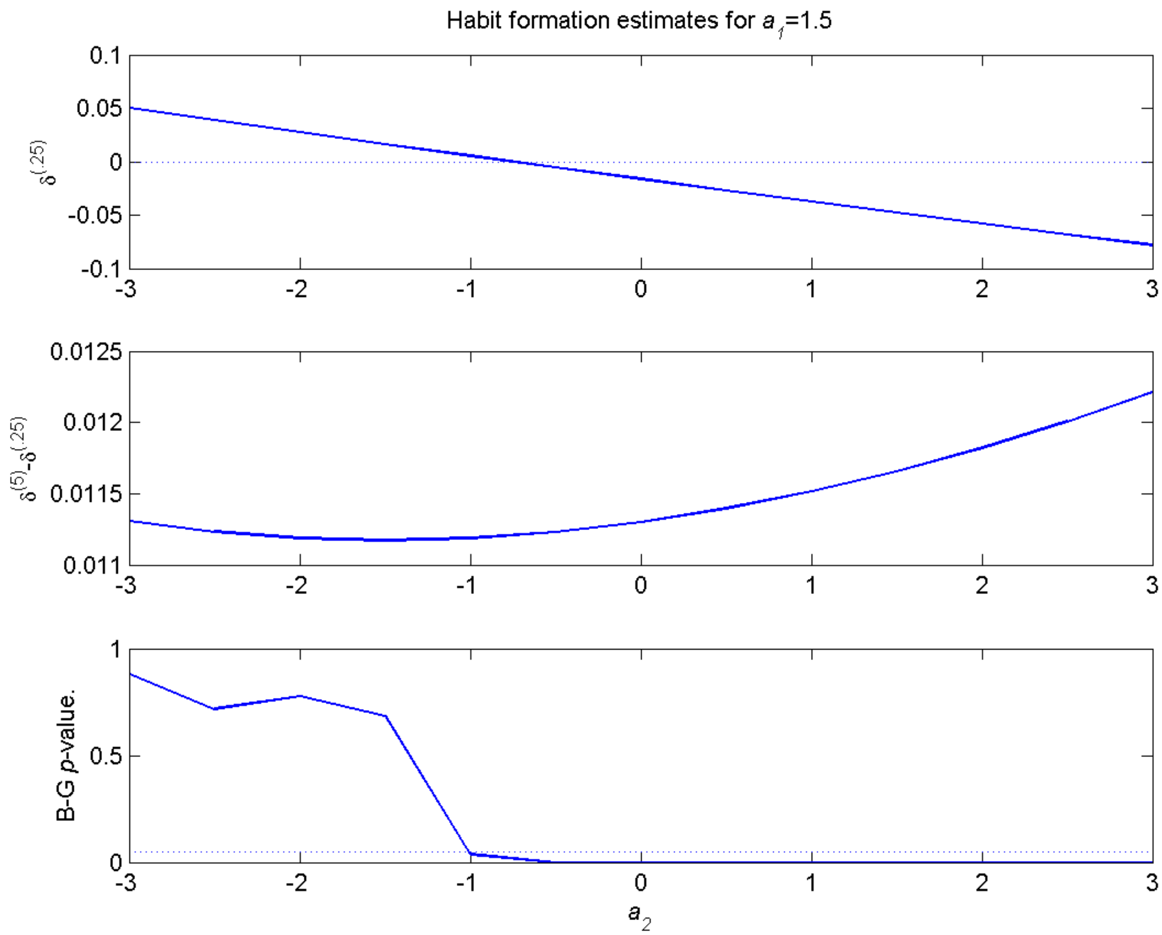
Note: Here we use the Bliss dataset that contains bond yields of longer maturities up to 10 years, but for the shorter sample period of 1970 to 1990. We again plot the discount premium $\delta^{(m)} - \delta^{(.25)}$ against the maturity. Since the estimates for the three values of α are very close (i.e., the three solid blue lines and the three dotted red lines), the reader is referred to the Online Appendix for the exact point estimates. Please refer to section 4.1 for details.

Figure 5: Alternative 95 Percent Confidence Intervals add "percent" after "95" below?



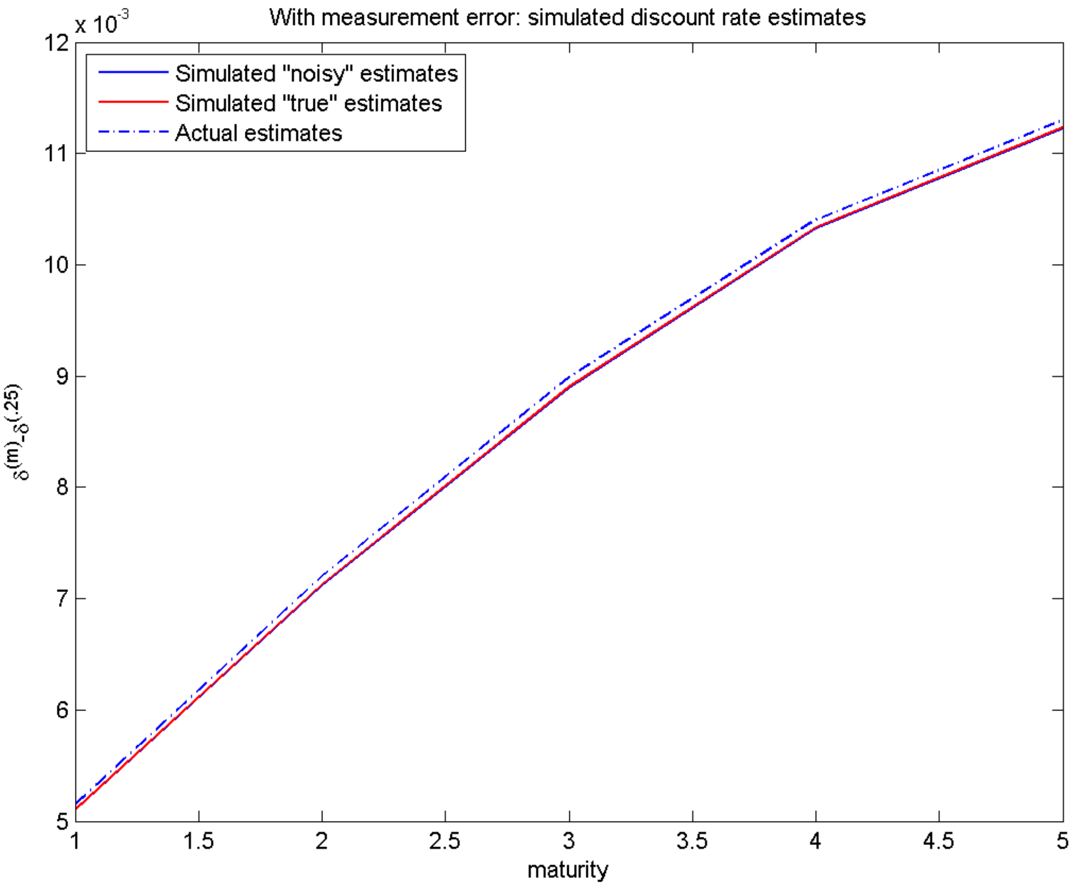
Note: We again plot the discount premium $\delta^{(m)} - \delta^{(.25)}$ against the maturity, but here we also provide confidence intervals from two parametric bootstraps. In the first case artificial data is generated from a residual from the estimated one-quarter Euler equation resampled with replacement (red lines). It assumes the errors in the estimated one-quarter Euler equation are serially uncorrelated. In the second bootstrap we allow serial correlation and resample from the estimated innovations, with ϵ_t as an ARMA(1,1) process (green lines). Please refer to section 4.2 for details.

Figure 6: Habit Formation



Note: Here we consider Abel's (1999) external habit formation model. On the x-axis we have different values of $a_2 = (1-\alpha)\boldsymbol{h_1}$ while fixing $a_1 = (1-\alpha)\boldsymbol{h_0} + \alpha$ at 1.5. In the first panel we plot the one-quarter discount rate. In the second panel we plot the discount premium for the maturity of 5 years. In the third panel we show the p -value for the Breusch–Godfrey test. A value above the dotted line of 0.05 means we do not reject the null of no serial correlation at the 5 percent level. Please refer to section 4.3 for details.

Figure 7: Allowing for Measurement Errors



Note: We again plot the discount premium $\delta^{(m)} - \delta^{(.25)}$ against the maturity. Here we conduct a Monte Carlo experiment in which log consumption follows a random walk with a shock with half the growth variance in the data (the "true" data). We then add to the simulated *level* of consumption an error scaled to the other half of the variance, giving simulated "noisy" consumption. We use the two sets of simulated data to estimate the discount premium. Please refer to section 4.5 for details.

Appendix

1. Data Source

Our sample is quarterly from 1954Q1 to 2007Q4.

Nominal yields: The one-quarter yield data are last-day-of-quarter observations of the 3-Month Treasury Bill (Secondary Market) from the FRED dataset compiled by the St. Louis Fed. The one- to 5-year zero-coupon yields are last-day-of-quarter observations from the Fama-Bliss Discount Bond Files of the U.S. Treasury database provided by the CRSP. (As an aside, while in principle alternative estimates could be made with real bonds, the history of short-term TIPS is still quite brief. See Gürkayank et al., Figure 1.)

Real consumption: Data on per capita real consumption of nondurable goods and services are from Table 7, *Selected Per Capita Product and Income Series in Current and Chained Dollars*, of the National Income and Product Accounts Table compiled by the Bureau of Economic Analysis. The series are in 2000 chained dollars.

Price level: To be consistent with the types of consumption we use, price level is defined by the weighted average of the price indexes of nondurable goods and the price index of service. The two indexes are from Table 1.1.4, Price Indexes for Gross Domestic Product, of the National Income and Product Accounts Table compiled by the Bureau of Economic Analysis. The weight is calculated using the two real consumption series described above. Specifically, on any quarter t, we calculate the weight as $w = \frac{C_{nondurables}}{C_{nondurables} + C_{services}}$, where C_x is the real per capita consumption of type x, then calculate the price index as $wP_{nondurables} + (1-w)P_{services}$.

Bliss yields: Bliss yields are last month of quarter, 1970-2000, and are supplied courtesy of Prof. Robert Bliss.

2. Epstein-Zin Estimation Procedure

We follow Piazzesi and Schneider (2007) and first estimate a state-space model for consumption growth and inflation, which we put together as $z_t = (\Delta c_t, \pi_t)'$:

$$z_t = \mu_z + x_{t-1} + e_t$$

$$x_t = \phi_x x_{t-1} + \phi_x K e_t$$

where $e_t \sim N(0, \Omega)$. The autoregressive matrix ϕ_x has four parameters, the gain matrix K has four parameters, and the covariance matrix Ω has three parameters. The model above becomes the simple first-order VAR when $K = I$. To see that, let $x_{t-1} = \phi_x(z_{t-1} - \mu_z)$:

$$z_t = \mu_z + \phi_x(z_{t-1} - \mu_z) + e_t$$

$$\phi_x(z_t - \mu_z) = \phi_x^2(z_{t-1} - \mu_z) + \phi_x e_t$$

The state-space model is a more general VARMA(1,1) process. We estimate the model by maximum likelihood by making use of the fact that $x_t = \phi_x x_{t-1} + \phi_x K(z_t - x_{t-1})$ and assuming $x_0 = 0$. To minimize the effect of the initial value x_0 we use the longer sample from 1947Q2 to 2007Q4; however, shortening the sample period does not affect the estimates significantly. When we are matching the yields below we are using the sample 1954Q1-2007Q4, the same sample we use in the paper.

As shown in Piazzesi and Schneider (2007), when the horizon is infinite the log real pricing kernel m under the Epstein-Zin preference can be written as:

$$m_{t+1} = \ln \beta - \Delta c_{t+1} - (\gamma - 1)(E_{t+1} - E_t)\sum_{i=0}^{\infty} \beta^{i+1} \Delta c_{t+1+i} - \frac{1}{2}(\gamma - 1)^2 var_t\left(\sum_{i=0}^{\infty} \beta^{i+1} E_{t+1}(\Delta c_{t+1+i})\right)$$

The parameter γ is the risk aversion parameter, and the model reduces to the CRRA case when $\gamma = 1$. The Epstein-Zin preference differs from the standard CRRA preference by allowing two terms related to the *temporal distribution of risk*: consumers fear downward revisions in consumption growth expectations and their uncertainty. The nominal pricing kernel is defined as $m_{t+1}^{\$} = m_{t+1} - \pi_{t+1}$.

Once we obtain the estimates from the state-space model, we need to choose the discount factor β and risk aversion parameter γ. Finally, we can make use of the nonarbitrage condition for nominal bond prices $P^{(n)}$:

$$P_t^{(n)} = E_t\left(P_{t+1}^{(n-1)} M_{t+1}^{\$}\right) = E_t\left(\Pi_{i=1}^{n} M_{t+i}^{\$}\right)$$

and calculate the implied bond yields $r_t^{(n)} = -\frac{1}{n} p_t^{(n)}$ recursively, using the definition of the log pricing kernel above.

How should the two parameters β and γ be chosen? In Piazzesi and Schneider (2007) the two parameters are chosen to match the means of the model one-quarter and five-year yields (the yields of the shortest and longest maturities in the data) with their corresponding sample values; the goal is to fit the two ends of the yield curve. Since our goal is to check for nonexponential discounting, we instead fix γ at some reasonable value and choose a discount factor β to match the model mean of *each* yield to its sample value. As a result, we obtain six different β factors for six different maturities, all with the same risk aversion parameter γ.

3. Extended Results

Here point estimates and Newey-West standard errors behind **Error! Reference source not found.** are presented. Estimates are by single-equation least squares.

α	$\widehat{\delta^{(.25)}}$	Standard error
0	0.016	0.003
0.1	0.014	0.003
0.2	0.012	0.003
0.3	0.010	0.003
0.4	0.007	0.003
0.5	0.005	0.003
0.6	0.003	0.003
0.7	0.001	0.003
0.8	-0.001	0.003
0.9	-0.004	0.004
1	-0.006	0.004
1.1	-0.008	0.004
1.2	-0.010	0.004
1.3	-0.012	0.004
1.4	-0.014	0.004
1.5	-0.016	0.004
1.6	-0.019	0.004
1.7	-0.021	0.004
1.8	-0.023	0.004
1.9	-0.025	0.004
2	-0.027	0.005
2.1	-0.029	0.005
2.2	-0.031	0.005
2.3	-0.033	0.005
2.4	-0.035	0.005
2.5	-0.037	0.005
2.6	-0.039	0.005
2.7	-0.042	0.005
2.8	-0.044	0.006
2.9	-0.046	0.006
3	-0.048	0.006

Error! Reference source not found. shows the discount rates and confidence intervals for $\alpha = 0.5, 1, \text{and} 3$. Here is the complete set of discount rates, the *p*-value testing for equality of the discount rates, and the *p*-value for testing the restrictions implicit in quasi-hyperbolic discounting.

α	$\delta^{(.25)}$	$\delta^{(1)}$	$\delta^{(2)}$	$\delta^{(3)}$	$\delta^{(4)}$	$\delta^{(5)}$	*p* (exp discounting)	*p* (quasi-hyper discounting)
0	0.016	0.021	0.023	0.025	0.027	0.028	0	0.0021
0.5	0.005	0.010	0.012	0.014	0.016	0.016	0	0.0249
1.0	-0.006	-0.001	0.002	0.003	0.005	0.006	0	0.2036
1.5	-0.016	-0.011	-0.009	-0.007	-0.006	-0.005	0.0001	0.601
2.0	-0.027	-0.022	-0.020	-0.018	-0.016	-0.016	0.0024	0.7493
2.5	-0.037	-0.032	-0.030	-0.028	-0.027	-0.026	0.0227	0.6382
3.0	-0.048	-0.042	-0.040	-0.038	-0.037	-0.036	0.0875	0.4507

Newey-West standard errors for the difference $\delta^{(m)} - \delta^{(.25)}$:

α	$\delta^{(1)} - \delta^{(.25)}$	$\delta^{(2)} - \delta^{(.25)}$	$\delta^{(3)} - \delta^{(.25)}$	$\delta^{(4)} - \delta^{(.25)}$	$\delta^{(5)} - \delta^{(.25)}$
0	0.001	0.002	0.002	0.003	0.003
0.5	0.001	0.002	0.002	0.002	0.002
1.0	0.001	0.002	0.002	0.002	0.002
1.5	0.001	0.002	0.002	0.003	0.003
2.0	0.002	0.003	0.003	0.003	0.003
2.5	0.002	0.003	0.003	0.004	0.004
3.0	0.002	0.004	0.004	0.004	0.005

Point estimates of the short discount rate from the habit formation model:

	a_1						
a_2	0.000	0.500	1.000	1.500	2.000	2.500	3.000
-3.000	0.085	0.073	0.062	0.050	0.039	0.028	0.017
-2.500	0.073	0.062	0.050	0.039	0.028	0.017	0.006
-2.000	0.062	0.050	0.039	0.028	0.016	0.006	-0.005
-1.500	0.050	0.039	0.027	0.016	0.005	-0.005	-0.016
-1.000	0.039	0.027	0.016	0.005	-0.006	-0.016	-0.027
-0.500	0.027	0.016	0.005	-0.006	-0.016	-0.027	-0.037
0.000	0.016	0.005	-0.006	-0.016	-0.027	-0.037	-0.048
0.500	0.005	-0.006	-0.016	-0.027	-0.037	-0.048	-0.058
1.000	-0.006	-0.016	-0.027	-0.038	-0.048	-0.058	-0.068
1.500	-0.016	-0.027	-0.038	-0.048	-0.058	-0.068	-0.078
2.000	-0.027	-0.038	-0.048	-0.058	-0.068	-0.078	-0.088
2.500	-0.037	-0.048	-0.058	-0.068	-0.078	-0.088	-0.098
3.000	-0.048	-0.058	-0.068	-0.078	-0.088	-0.098	-0.107

Breusch-Godfrey *p*-values from the habit formation model:

	a_1						
a_2	0.000	0.500	1.000	1.500	2.000	2.500	3.000
-3.000	0.000	0.000	0.044	0.883	0.051	0.001	0.000
-2.500	0.000	0.000	0.054	0.718	0.037	0.002	0.001
-2.000	0.000	0.000	0.043	0.776	0.077	0.015	0.013
-1.500	0.000	0.000	0.010	0.685	0.475	0.302	0.366
-1.000	0.000	0.000	0.000	0.038	0.210	0.287	0.243
-0.500	0.000	0.000	0.000	0.000	0.000	0.000	0.001
0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.500	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1.500	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2.500	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Point estimates of $\delta^{(m)}$ using the Bliss data:

	α						
m	*0.000*	*0.500*	*1.000*	*1.500*	*2.000*	*2.500*	*3.000*
0.250	0.019	0.008	-0.002	-0.012	-0.022	-0.032	-0.042
0.500	0.022	0.011	0.001	-0.010	-0.020	-0.030	-0.040
0.750	0.023	0.013	0.002	-0.008	-0.018	-0.028	-0.038
1.000	0.024	0.014	0.003	-0.007	-0.017	-0.027	-0.036
1.250	0.025	0.015	0.005	-0.006	-0.016	-0.026	-0.035
1.500	0.026	0.016	0.005	-0.005	-0.015	-0.025	-0.034
1.750	0.027	0.017	0.006	-0.004	-0.014	-0.024	-0.034
2.000	0.027	0.017	0.006	-0.004	-0.014	-0.024	-0.034
2.500	0.029	0.018	0.008	-0.002	-0.012	-0.022	-0.032
3.000	0.030	0.019	0.009	-0.001	-0.011	-0.021	-0.030
4.000	0.032	0.022	0.012	0.002	-0.008	-0.018	-0.028
5.000	0.035	0.025	0.014	0.004	-0.006	-0.015	-0.025
6.000	0.038	0.028	0.017	0.007	-0.003	-0.012	-0.022
7.000	0.039	0.029	0.019	0.009	-0.001	-0.010	-0.020
8.000	0.041	0.031	0.021	0.011	0.002	-0.008	-0.017
9.000	0.043	0.033	0.023	0.013	0.004	-0.006	-0.016
10.000	0.044	0.034	0.024	0.014	0.005	-0.005	-0.014

As a robustness check, we split the sample into 1954-1978 and 1979-2002 subsamples. Subsample statistics are given below, along with the analogs of the earlier calculations absent uncertainty. Very little changes, though the effects are larger in the second part of the sample.

Subsample	$\bar{r}^{(.25)}$	$\bar{r}^{(5)}$	$\bar{g}$	$\bar{\pi}$	$var(g^{(.25)})$	$var(\pi^{(.25)})$	$cov(g^{(.25)}, \pi^{(.25)})$	max α	$\delta^{(5)} - \delta^{(.25)}$
1954-1978	4.27×10^{-2}	5.11×10^{-2}	0.0250	0.0356	2.25×10^{-5}	6.03×10^{-5}	-1.27×10^{-5}	0.247	0.0084
1979-2002	6.50×10^{-2}	7.83×10^{-2}	0.0184	0.0352	1.67×10^{-5}	4.51×10^{-5}	-8.17×10^{-6}	1.62	0.0133