

Euler Equation Errors*

Martin Lettau

New York University, CEPR, NBER

Sydney C. Ludvigson

New York University and NBER

PRELIMINARY

Comments Welcome

First draft: September 1, 2004

This draft: February 5, 2005

*Lettau: Department of Finance, Stern School of Business, New York University, 44 West Fourth Street, New York, NY 10012-1126; Email: mlettau@stern.nyu.edu, Tel: (212) 998-0378; Fax: (212) 995-4233; <http://www.stern.nyu.edu/~mlettau>. Ludvigson: Department of Economics, New York University, 269 Mercer Street, 7th Floor, New York, NY 10003; Email: sydney.ludvigson@nyu.edu, Tel: (212) 998-8927; Fax: (212) 995-4186; <http://www.econ.nyu.edu/user/ludvigsons/>. Ludvigson acknowledges financial support from the Alfred P. Sloan Foundation and the CV Starr Center at NYU. We are grateful to Dave Backus, John Y. Campbell, George Constantinides, Mark Gertler, Stephen Gordon, Fatih Guvenen, Stijn Van Nieuwerburgh, Martin Weitzman and seminar participants at New York University, Université Laval, and West Virginia University for helpful comments. We Thank Jack Favalukis for excellent research assistance. Any errors or omissions are the responsibility of the authors.

Euler Equation Errors

Abstract

Among the most important pieces of empirical evidence against the standard representative-agent, consumption-based asset pricing paradigm are the formidable unconditional Euler equation errors the model produces for a broad stock market index return and short-term interest rate. Unconditional Euler equation errors are also large for a broader cross-section of returns. Here we ask whether calibrated leading asset pricing models—specifically developed to address empirical puzzles associated with the standard paradigm—explain these empirical facts. We find that, in many cases, they do not. We present several results. First, we show that if the true pricing kernel that sets the unconditional Euler equation errors to zero is jointly lognormally distributed with aggregate consumption and returns, then values for the subjective discount factor and relative risk aversion can always be found for which the standard model generates identical unconditional asset pricing implications for two asset returns, a risky and risk-free asset. Second, we show, using simulated data from several leading asset pricing frameworks, that many economic models share this property even though in those models the pricing kernel, returns, and consumption are not jointly lognormally distributed. Third, in contrast to the above results, we provide an example of a limited participation/incomplete markets model that is broadly consistent with these empirical facts.

JEL: G12, G10.

1 Introduction

Among the most important pieces of empirical evidence against the standard representative-agent, consumption-based asset pricing paradigm are the formidable unconditional Euler equation errors the model produces for a broad stock market index return and short-term interest rate. Our definition of the *standard* model assumes that agents have unrestricted access to financial markets, that assets can be priced using the Euler equations of a representative-consumer maximizing the discounted value of power utility functions, and that the pricing kernel M , or stochastic discount factor, is equal to the marginal rate of substitution in consumption. This model takes the form

$$E[M_t R_t] = 1, \quad M_t = \delta (C_t/C_{t-1})^{-\gamma}, \quad (1)$$

where R_t denotes the gross return on any tradable asset, C_t is per capita aggregate consumption, γ is the coefficient of relative risk-aversion and δ is a subjective time-discount factor. The average Euler equation errors are also large for a broader cross-section of returns that includes size and book-market sorted portfolio returns.

We argue that these Euler equation errors constitute a puzzle for the standard consumption-based asset pricing model that is at least as damning as other, more well known, conundrums that have received far more attention, such as the equity premium puzzle, risk-free rate puzzle, and time-series predictability of excess stock market returns. We employ these empirical facts to evaluate leading asset pricing models that have been specifically developed to address puzzles generated by the standard paradigm (1). If leading asset pricing models are true, then in these models using (1) to price assets should generate large unconditional asset pricing errors, as in the data. The underlying assumption in each of these leading models is that, by discarding the standard pricing kernel in favor of the true kernel implied by the model, an econometrician would be better able to model asset pricing data.

In this paper we show that this is not always the case. Often, in leading asset pricing models, a standard representative-agent “pricing kernel” based on (1) can be found that has virtually identical unconditional pricing implications for specific asset returns, such as a risky asset and risk-free asset, or for a larger cross-section of risky asset returns. Thus, an econometrician who observed data generated from any number of these leading models would fail to reject the standard consumption-based model in tests of its unconditional moment restrictions, let alone replicate the sizable unconditional Euler equation errors found when fitting historical data to (1).

The literature has already demonstrated that it is possible, in principle, to explain any observed behavior of per capita aggregate consumption and asset returns, by appealing to incomplete consumption insurance. Constantinides and Duffie (1996) prove a set of theoretical

propositions showing that any observed joint process of aggregate consumption and returns can be an equilibrium outcome if the second moments of the cross-sectional distribution of consumption growth and asset returns covary in the right way. Krebs (2004) shows that any observed joint process of aggregate consumption and asset returns can be rationalized if all asseholders are subject to sufficiently extreme idiosyncratic events with very small probability of occurrence. In this paper we move away from theoretical propositions and ask whether particular calibrated economies of leading asset pricing models are quantitatively capable of matching the large unconditional Euler equation errors generated by the standard consumption-based model when fitted to data. This is important because it remains unclear whether plausibly calibrated models built on primitives of tastes, technology, and underlying shocks can in practice generate the joint behavior we observe in the data.

Our analysis uses simulated data from several leading economic models designed address empirical failures of the standard model (1). These models include the representative-agent external habit-persistence paradigms of (i) Campbell and Cochrane (1999) and (ii) Menzly, Santos, and Veronesi (2004), (iii) the representative-agent long-run risks model based on recursive preferences of Bansal and Yaron (2004), and (iv) the limited participation model of Guvenen (2003). Each is an explicitly parameterized economic model calibrated to accord with the data in plausible ways, and each has proven remarkably successful in explaining a range of asset pricing phenomena. In addition to these models, we dig more deeply into the aggregate Euler equation implications of simple asset pricing models with limited participation/incomplete markets, in which asseholder consumption is permitted to behave quite differently from per capita aggregate consumption.

Our focus on Euler equations is intentional, since they represent the set of theoretical restrictions from which all asset pricing implications follow. Formal econometric tests of *conditional* Euler equations using aggregate consumption data lead to rejections of the standard representative-agent, consumption-based asset pricing model, even when no bounds are placed on the coefficient of relative risk aversion or the rate of time preference (Hansen and Singleton (1982); Ferson and Constantinides (1991); Hansen and Jagannathan (1991)). Similarly, we show here that the quarterly pricing errors for the *unconditional* Euler equations associated with an aggregate equity return and a short-term Treasury bill rate are large when fitting aggregate data to (1), even when the coefficient of relative risk aversion or the rate of time preference are left unrestricted and chosen to minimize those errors. For larger cross-sections of returns the results are similar. These empirical results place additional testable restrictions on asset pricing models: not only must such models have zero pricing errors when the pricing kernel is correctly specified according to the model, they must also produce large pricing errors when the pricing kernel is incorrectly specified using power utility and aggregate consumption, even when δ and γ are chosen to minimize those errors.

Our main findings are as follows:

First, we show that if the true pricing kernel that sets Euler equation errors to zero is jointly lognormally distributed with aggregate consumption and returns, then values for the discount factor and relative risk aversion can be always be found for which the standard model generates identical unconditional asset pricing implications for two asset returns, a risky and risk-free asset. This property implies that such models will not be capable of explaining the empirical facts discussed above, namely the large Euler equation errors found when asset return data are fitted to (1). We illustrate these results in an incomplete markets/limited participation setting.

Second, using simulated data from each of the leading asset pricing models mentioned above, we show that many economic models share this property even though in these models the pricing kernel, returns, and consumption are not jointly lognormally distributed. Some of the models we study *can* explain why an econometrician obtains implausibly high estimates of δ and γ when freely fitting aggregate data to (1). But, they cannot explain the large unconditional Euler equation errors associated with such estimates. The asset pricing models we consider counterfactually imply that values for the subjective discount factor and risk aversion can be found for which (1) satisfies the unconditional Euler equation restrictions just as well as the true pricing kernel.

Third, in contrast to the above results, we provide one example in an incomplete markets/limited participation setting of a model that can roughly replicate the empirical facts, if the joint distribution of aggregate consumption, individual asseholder consumption, and stock returns takes a particular form of deviations from normality. But we also find—within this broad class of distributions we consider—that many non-normal distribution specifications do not explain the sizeable Euler equation errors generated by the standard consumption-based asset pricing model (1). Similar findings hold for the average Euler equation errors over a larger cross-section of asset returns.

We emphasize that this paper is not a criticism of work that investigates the asset pricing implications of models with preferences or market structures that differ from the standard consumption-based model. Indeed, we view our paper as a compliment to the existing literature because it provides a different perspective on whether such models are capable of fully rationalizing the joint behavior of asset prices and aggregate quantities. We also add to the literature by outlining the econometric consequences, for estimation and testing of unconditional Euler equations, of fitting the standard pricing kernel (1) to data when the true pricing kernel that generated the data is derived from some other model. Finally, we stress that our results do not imply that no model can be made consistent with the testable restrictions we focus on here—we present an incomplete markets example to the contrary. Our point is that many models, including those at the forefront of asset pricing theory, do

not satisfy these testable restrictions.

The rest of this paper is organized as follows. The next section lays out the empirical Euler equation facts using post-war U.S. data on per capita aggregate consumption and returns. Section 3 studies the implications of various economic theories for the same Euler equation errors we measure in the data. We begin with a simple example in which the true pricing kernel is jointly lognormally distributed with aggregate consumption growth and asset returns. Next, we investigate the extent to which leading asset pricing models, calibrated to accord with the U.S. data, are capable of explaining the empirical facts. Here we focus both on the case of a single risky and a risk-free asset return and on models that exploit a larger cross-section of risky returns. Our main findings are shown to be robust to time-aggregation of aggregate consumption data, and to the introduction of limited participation in the representative-agent models. Finally, we provide one example of a simple incomplete markets/limited participation model that can roughly replicate the empirical Euler equation errors from historical data. Section 4 concludes.

2 Euler Equation Errors: The Facts

In this section we document empirical facts of the standard consumption-based asset pricing model that pertain to unconditional Euler equations using aggregate consumption.

The standard consumption based model, as defined above, has the following characteristics. There exists representative-consumer with constant relative risk aversion (CRRA) preferences over consumption given by

$$U = E \left\{ \sum_{t=0}^{\infty} \delta^t \frac{C_t^{1-\gamma} - 1}{1-\gamma} \right\}, \quad \gamma > 0, \quad (2)$$

where C_t is per capita aggregate consumption, δ is a subjective time-discount factor, and γ is the coefficient of relative risk aversion. At each date, agents maximize (2) subject to an accumulation equation for wealth that links assets today to returns today and savings yesterday. Agents have unrestricted access to financial markets; they face no borrowing or short-sales constraints. The asset pricing model comes from the first-order conditions for optimal consumption choice, which place restrictions on the joint distribution of the intertemporal marginal rate of substitution in consumption, given by $M_{t+1} \equiv \delta (C_{t+1}/C_t)^{-\gamma}$, and asset returns. The first-order conditions imply that for any traded asset indexed by j , with a gross return at time $t + 1$ of R_{t+1}^j , the following set of Euler equation holds:

$$E_t [M_{t+1} R_{t+1}^j] = 1. \quad (3)$$

Here E_t is the conditional expectation operator, conditional on time t information. The marginal rate of substitution in consumption, M_{t+1} , is the stochastic discount factor, or pricing kernel.

By the law of iterated expectations, equation (3) also implies the following set of unconditional Euler equations that must hold for any traded asset indexed by j :

$$E [M_{t+1}R_{t+1}^j] = 1, \quad (4)$$

where E is the unconditional expectation operator.

We refer to the difference between the left-hand-side of (4) and unity as the *unconditional Euler equation error*, or alternatively the *pricing error*, associated with the j th asset return:

$$\text{Pricing error of asset } j = E [\delta (C_t/C_{t-1})^{-\gamma} R_{t+1}^j] - 1.$$

If the standard model is true then these errors should be zero for any traded asset, given some values of the parameters δ and γ .

We focus much of our attention on the unconditional Euler equation errors associated with two asset returns: a broad stock market index return (measured as the the CRSP value-weighted price index return and denoted R_t^s) and short-term “riskless” interest rate (measured as the three-month Treasury bill rate and denoted R_t^f). We maintain this focus for two reasons. First, the standard model’s inability to explain properties of just these two returns has been central to the development of a consensus that the model is seriously flawed. Second, all asset pricing models seek to match the empirical properties of these two returns and most generate no implications for larger cross-sections of securities because the cash flow properties of these securities are not modeled. Nevertheless, we also consider the Euler equation errors associated with a broader cross-section of returns, including portfolio returns formed on the basis of size and book-market ratios. Notice that estimation of (4) for the two-asset case collapses to an exercise in solving two nonlinear equations. In this case, the vector of sample pricing errors contains no nondegenerate random variables and thus the pricing errors are not affected by sampling error.

There are several ways to present the pricing errors implied by the standard consumption-based model for these two asset returns. One is to combine the separate Euler equations for the stock market return and Treasury bill rate into a single Euler equation for the excess return that is a function of only the risk aversion parameter γ . From (4) we have

$$E \left[\left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \left(R_{t+1}^s - R_{t+1}^f \right) \right] = 0. \quad (5)$$

The empirical pricing error for the excess return is given by an estimate of the left-hand-side of (5). Figure 1 plots this error over a range of values of γ , where the error is computed as

the sample mean of the expression in square brackets in (5). The estimation uses quarterly, per capita data on nondurables and services expenditures measured in 1996 dollars as a measure of consumption C_t , in addition to the return data mentioned above.¹ Returns are deflated by the implicit price deflator corresponding to the measure of consumption C_t . The data span the period from the first quarter of 1951 to the fourth quarter of 2002. A detailed description is provided in the Appendix.

Figure 1 shows that the pricing error (5) cannot be driven to zero, or indeed even to a small number, for any value of γ . The lowest pricing error is almost 4% per annum, which occurs at $\gamma = 117$. This pricing error is almost *half* of the average annual CRSP stock return and *four* times the average annual Treasury-bill rate. At other values of γ this error rises precipitously and reaches several times the average annual stock market return when γ is outside the ranges displayed in Figure 1. Thus, there is no value of γ that sets the pricing error (5) to zero.²

Another way to present the pricing errors implied by the standard model for the stock market return and Treasury bill rate is to choose values for δ and γ that make the estimated Euler equation errors for each asset return as close as possible to the theoretical Euler equation errors of zero:

$$\begin{aligned} E \left[\delta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} R_{t+1}^s \right] - 1 &= 0 \\ E \left[\delta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} R_{t+1}^f \right] - 1 &= 0. \end{aligned}$$

This amounts to solving two (nonlinear) equations in two unknowns, δ and γ . To do so, we choose parameters for time preference δ and risk aversion γ to minimize a weighted sum of squared pricing errors, an application of exactly identified Generalized Method of Moments (GMM, Hansen (1982)):

$$\min_{\delta, \gamma} g_T(\gamma, \delta) = \omega_s E \left[\delta (C_t/C_{t-1})^{-\gamma} R_t^s - 1 \right]^2 + \omega_f E \left[\delta (C_t/C_{t-1})^{-\gamma} R_t^f - 1 \right]^2,$$

for some weights ω_s and ω_f . Let δ_c and γ_c denote the $\arg \min g_T(\gamma, \delta)$, the values of δ and γ that minimize the scalar $g_T(\gamma, \delta)$.

Panel A of Table 1 shows that when the risk aversion and time preference parameters are chosen to minimize an equally weighted sum of squared pricing errors ($\omega_s = \omega_f = 1$), the pricing errors for the stock market return and the short-term Treasury bill rate are split evenly, with each error close to 2.7% per annum for the stock market return and close to

¹We exclude shoes and clothing expenditure from this series since they are partly durable.

²Note that (5) is a nonlinear function of γ . Thus, there is not necessarily a solution.

-2.7% for the Treasury bill. As before, this magnitude is large: it is about one-third of the average annual stock market return, and about three times as large as the average annual Treasury bill return. The last column shows that the square root of the average squared pricing error is 60% of the cross-sectional mean return for these two assets. Naturally, when we place 100 times more weight on the Euler equation associated with the stock market return in this minimization ($\omega_s = 100, \omega_f = 1$), the pricing error associated with the stock return can be made close to zero, but then the error associated with the Treasury bill explodes to -5.4% per annum. Conversely, when we place 100 times more weight on the Euler equation associated with the Treasury bill rate ($\omega_s = 1, \omega_f = 100$), the pricing error associated with the Treasury bill can be made close to zero, but the error associated with the stock return rises to 5.4% per annum. Thus, there are no values of δ and γ that set the pricing errors to zero. Regardless of the particular weighting scheme, δ_c and γ_c (which are left unrestricted) are close to 1.4 and 90, respectively.

Figure 2 displays the (negative of) the square root of the averaged squared pricing errors over a range of values for δ and γ . The figure is presented this way for ease of readability. The pricing errors are obviously smallest at the point estimates for δ and γ , but the figure shows that they are large over a wide range of values for these parameters. Finally, Figure 3 shows the contour plots for the output displayed in Figure 2. The figures show that there is no combination of parameter values for δ and γ for which the pricing errors of both asset returns can be set to zero, or even small in magnitude.

Why are the pricing errors for the stock return and Treasury-bill rate so large? Panel B of Table 1 provides a partial answer: a large part (but not all) of the unconditional Euler equation errors generated by this model are associated with recessions, periods in which per capita aggregate consumption growth is steeply negative. If we remove the data points associated with the smallest six observations on consumption growth, the equally weighted pricing errors are 0.7% for the stock market return and -0.7% for the Treasury-bill rate, a 74% reduction. Table 2 identifies these six observations as they are located throughout the sample. Each occur in the depths of recessions in the 1950s, 1970s, early 1960s, 1980s and 1990s, as identified by the National Bureau of Economic Research. In these periods, aggregate per capita consumption growth is steeply negative but the aggregate stock return and Treasury-bill rate is, more often than not, steeply positive. Since the product of the marginal rate of substitution and the gross asset return must be unity on average, such negative comovement (positive comovement between M_t and returns) contributes to large pricing error. One can also reduce the (equally-weighted) pricing errors by using annual returns and year-over-year consumption growth (fourth quarter over fourth quarter).³ This

³Jagannathan and Wang (2004) study the ability of the standard model to explain a large cross-section

procedure averages out the worst quarters for consumption growth instead of removing them. Either procedure eliminates much of the cyclical variation in consumption. To see this, note that on a quarterly basis the largest declines in consumption are about six times as large at an annual rate as those on a year-over-year basis.

Of course, these quarterly episodes are not outliers to be ignored, but significant economic events to be explained. Indeed, we argue that such Euler equation errors—driven by periods of important economic change—are among the most damning pieces of evidence against the standard model. An important question is why the standard model performs so poorly in recessions relative to other times.

The pricing error of the Euler equation associated with the stock market return is always positive implying a positive “alpha” in the expected return-beta representation. This says that unconditional risk premia are too high to be explained by the stock market’s covariance with the marginal rate of substitution of aggregate consumption, a familiar result from the equity premium literature.⁴ The high alphas generated by the standard consumption-based model constitute one of the most remarked-upon failures in the history of asset pricing theory.

What about larger cross-sections of returns? Panel B of Table 1 reports the pricing errors for a cross-section of eight asset returns. The asset returns include as before the Treasury bill rate and CRSP stock market return, in addition to six value-weighted portfolio returns of common stock sorted into two size (market equity) quantiles and three book value-market value quantiles. These returns were taken from Professor Kenneth French’s Dartmouth web page. The latter six portfolios are created from stocks traded on the NYSE, AMEX, and NASDAQ, as detailed on Kenneth French’s web page. We use equity returns on

of asset returns using fourth quarter over fourth quarter consumption growth and annual asset returns. They find more support for the model when year-over-year growth rates are restricted to the fourth quarter.

⁴The expected return-beta representation is derived from the Euler equation

$$E [M_t R_t^s] - 1 = e,$$

where e denotes the pricing error associated with the Euler equation. This equation can be rearranged to yield

$$E (R_t^s) - 1/E (M_t) = \frac{-\text{Cov} (M_t, R_t^s) \text{Var} (M_t)}{\text{Var} (M_t) E (M_t)} + \frac{e}{E (M_t)},$$

where $1/E (M_t)$ is interpreted as the risk-free rate in models for which this rate is presumed constant. Rewrite the above as

$$E (R_t^s) - 1/E (M_t) = \beta \lambda + \alpha,$$

where the left-hand-side is the risk-premium on the stock return, $\beta \equiv \frac{-\text{Cov}(M_t, R_t^s)}{\text{Var}(M_t)}$ is the “beta” for the stochastic discount factor M_t , or quantity of risk, $\lambda \equiv \frac{\text{Var}(M_t)}{E(M_t)}$ is the price of risk associated with M_t , and $\alpha \equiv \frac{e}{E(M_t)}$, is the “alpha” associated with the market return, that is the part of the risk-premium that cannot be explained by its beta risk.

size and book-to-market sorted portfolios because Fama and French (1992) show that these two characteristics provide a “simple and powerful characterization” of the cross-section of average stock returns, and absorb the roles of leverage, earnings-to-price ratio and many other factors governing cross-sectional variation in average stock returns. For this set of asset returns parameters γ and δ are chosen to minimize the quadratic form $g_T(\gamma, \delta) \equiv \mathbf{w}'_T(\gamma, \delta) \mathbf{W} \mathbf{w}_T(\gamma, \delta)$, where $\mathbf{w}_T(\gamma, \delta)$ is the (8×1) vector of average pricing errors for each asset (i.e., $w_{jT}(\gamma, \delta) = \frac{1}{T} \sum_{t=1}^T \delta \left(\frac{C_t}{C_{t-1}} \right)^{-\gamma} R_t^j - 1$ for $j = 1, \dots, 8$) and \mathbf{W} is the 8×8 identity matrix. We use the identity weighting matrix because we are interested in the average pricing errors for these particular returns, which Fama and French show are based on economically interesting characteristics and deliver a wide spread in cross-sectional average returns. Using the optimal GMM weighting matrix (for example) would require us to minimize the pricing errors for re-weighted portfolios of the original test assets.

Panel B shows that the square root of the average squared pricing error (RMSE) for the eight asset returns is large: over three% on an annual basis, or over 35% of the cross-sectional average mean return. Eliminating the lowest six consumption growth rate periods reduces the average pricing errors, but they remain large, around 2% on an annual basis.

How much might sampling error alone contribute to the estimated Euler equation errors for the stock return and Treasury bill rate? In principle, this question can be addressed in exactly identified systems by conducting a block bootstrap simulation of the raw data. This approach is inappropriate for the application here, however, because such a procedure would effectively treat the low consumption growth periods in our sample as outliers, in the sense that a nontrivial fraction of the simulated samples would exclude those observations. But as we have argued above, these episodes of low or negative consumption growth—the hallmark of recessions—are not outliers to be ignored, but significant economic events to be explained.

A more appropriate approach to this question is to ask, given sampling error, how likely is it that we would observe the pricing errors we observe under the null hypothesis that the standard model is true and the Euler equations are exactly satisfied in population? Models that postulate joint lognormality for consumption and asset returns are null models of this form, since in this case values for δ and γ always exist for which the population Euler equations of any two asset returns are exactly satisfied. Consequently, only sampling error in the estimated Euler equations could cause non-zero pricing errors. It is therefore natural to begin by assessing whether joint lognormality is a plausible description of our consumption and return data, once we take into account sampling error. We do so by performing formal statistical tests of lognormality in our data. Table 3 presents the results of normality tests, based on estimates of both univariate and multivariate skewness and kurtosis for $\log(C_t/C_{t+1}) \equiv \Delta c_t$, $\log(R_{st}) \equiv r_{st}$, and $\log(R_{ft}) \equiv r_{ft}$.

There is strong evidence against normality from both univariate and multivariate tests. We reject that skewness is zero at the one% or better level for Δc_t , and r_{st} , for every pair of the three variables Δc_t , r_{st} , and r_{ft} , and for the triple $(\Delta c_t, r_{st}, r_{ft})$. Also at better than the one% level, we may reject the null hypothesis that the kurtosis measures for any three of Δc_t , r_{st} , and r_{ft} are equal to those of a univariate normal distribution, and that the kurtosis measure for any pair of these variable or for the triple $(\Delta c_t, r_{st}, r_{ft})$ is equal to that of a multivariate normal distribution. The same picture emerges from examining quantile-quantile plots (*QQ* plots) for Δc_t , r_{st} , and r_{ft} and for the three variables jointly, given in Figure 4. This figure plots the sample quantiles for the data against those that would arise under the null of joint lognormality. Pointwise standard error bands are computed by simulating from the appropriate distribution with length equal to the size of our data set. The figure shows that all three log variables have some quantiles that lie outside the standard error bands for univariate normality. But the most dramatic departure from normality is displayed in the multivariate *QQ* plot for the joint distribution of $(\Delta c_t, r_{st}, r_{ft})$. In this case, a vast range of quantiles lie far outside the standard error bands for joint normality.

We address the issue of sampling error in our application from another angle. Suppose that the data were generated by the standard CRRA representative-agent model, with returns and consumption jointly lognormally distributed. How likely is it that we would find results like those reported in Table 1, in a sample of the size we have? Again, in this case population Euler equation errors are identically zero, so only sampling error in the estimated Euler equations can cause non-zero pricing errors. It is straightforward to address this question in a simple model where $\Delta \ln C_{t+1} \sim i.i.d.N(\mu, \sigma^2)$, and preferences are of the CRRA form with (for example) $\delta = 0.99$ and $\gamma = 2$. Since the log difference in consumption is i.i.d. and normally distributed, the return to a risky asset that pays consumption, C_t , as its dividend is also normally distributed, as is risk-free rate, $R_t^f \equiv 1/E_t[M_{t+1}]$. The equilibrium returns have an analytical solution in this case, and can be solved from the (exactly satisfied) Euler equations. Using this model, we simulate 1000 artificial samples of consumption data equal to the size our quarterly data set, with μ and σ set to match their respective sample estimates. Using the analytical solutions for returns we use the simulated data for consumption growth to obtain corresponding simulated data for returns. Finally, we use these simulated data to solve for the values of δ and γ that minimize the empirical Euler equation errors for the risky and risk-free asset return and store the absolute value of those errors. The 95% centered confidence for these errors, in percent annum, is found to be $(9.5 \cdot 10^{-11}, 7.0 \cdot 10^{-9})$ for the risky return and $(1.3 \cdot 10^{-10}, 6.5 \cdot 10^{-9})$ for the risk-free return. This reinforces the conclusion from the normality tests above, namely that sampling error alone is unlikely to account for the findings reported in Figures 1 and 2 and Table 1.

3 Euler Equation Errors: The Theories

How capable are leading asset pricing theories of correctly modeling the asset pricing phenomena described above? In this section, we address this question by considering a number of distinct asset pricing models. First we show that any model whose pricing kernel is jointly lognormally distributed with aggregate consumption and returns will fail to explain these features of the data, since in this case values for δ and γ can always be found for which the standard model has the same unconditional asset pricing implications as the true kernel for a risky and risk-free asset return. We illustrate this in a limited participation/incomplete markets setting. Next we evaluate the Euler equation errors generated by leading asset pricing models. As mentioned, these include the external habit-formation models of Campbell and Cochrane (1999) and Menzly, Santos, and Veronesi (2004), the long-run risks model of Bansal and Yaron (2004), and the limited participation model of Guvenen (2003). Finally, we present a number of additional results for simple limited participation/incomplete markets models in which assetholder consumption, aggregate consumption and asset returns are not jointly lognormally distributed.

3.1 Joint Lognormality: An Illustration using Limited Participation/Incomplete Markets Models

The common defining feature of limited participation and incomplete markets models is the presumption that it is not aggregate per capita consumption that is required to explain asset returns, but rather the consumption of some subset of the aggregate who are marginal asset holders. With limited stock market participation, the set of Euler equations of *stockholder* consumption imply that a representative stockholder's marginal rate of substitution is a valid stochastic discount factor and can be used to explain asset returns, but no Euler equations utilizing per capita aggregate consumption can be used to explain asset returns. Similarly, if incomplete consumption insurance means that heterogeneous consumers cannot equalize their marginal rate of substitution state-by-state, the set of Euler equations of *household* consumption imply that any household's marginal rate of substitution is a valid stochastic discount factor and can be used to explain asset returns, but no Euler equations utilizing per capita aggregate consumption can be used to explain asset returns. In other respects, these models are identical to the standard one: preferences are of the same von Neumann-Morgenstern form and assetholders face no frictions in accessing financial markets, as in (1). Thus, an econometrician who unwittingly attempted to fit data to (1) would be erring merely in using per capita aggregate consumption in the pricing kernel in place of stockholder or

individual assetholder consumption.⁵

To evaluate the unconditional pricing errors in models with limited stock market participation or incomplete markets, we must take a stand on the joint distribution of aggregate consumption, stockholder/individual consumption, and asset returns. As a benchmark case, we assume these to be jointly lognormally distributed. Later we consider asset pricing models in which the joint distribution is permitted to deviate from lognormality. In order to isolate the implications of these features for asset pricing, we keep these models standard in other respects; for example agents have standard time-separable preferences and unrestricted access to financial markets.

In analogy to the empirical investigation, consider the Euler equation errors for two assets with a gross returns denoted R_t^s and R_t^f . The first is a risky asset, for example a stock market return, and the second is a risk-free return. The next section will consider larger cross-sections of returns.

Denote the marginal rate of substitution (MRS) of an individual asset-holder as

$$M_{t+1}^i \equiv \delta \left(\frac{C_{t+1}^i}{C_t^i} \right)^{-\gamma}, \quad (6)$$

where C_t^i is the consumption of assetholder i , δ is the subjective time discount factor of this assetholder, and γ is the coefficient of relative risk aversion. If agents have unrestricted access to financial markets, then M_t^i correctly prices the two asset returns R_{t+1}^s and R_{t+1}^f , implying that

$$E_t [M_{t+1}^i R_{t+1}^s] = E_t [M_{t+1}^i R_{t+1}^f] = 1. \quad (7)$$

We focus on unconditional implications of these models,

$$E [M_{t+1}^i R_{t+1}^s] = E [M_{t+1}^i R_{t+1}^f] = 1.$$

We can interpret the MRS, M_{t+1}^i , either as that of a representative stockholder in a limited participation setting (C_t^i is then the consumption of a representative assetholder), or as that of an individual assetholder in an incomplete markets setting (C_t^i is the consumption of any marginal assetholder, e.g., Constantinides and Duffie (1996)).

Now denote the misspecified ‘‘MRS’’ for some parameters δ_c and γ_c , that would be computed if an econometrician erroneously used per capita aggregate consumption, C_t in place of C_t^i

$$M_t^c \equiv \delta_c \left(\frac{C_t}{C_{t-1}} \right)^{-\gamma_c}. \quad (8)$$

⁵Alternatively, one can interpret the example in this section as an illustration of the influence of measurement error on empirically observed pricing errors. In this case, stockholder consumption corresponds to correctly measured consumption for which the model holds exactly, and aggregate consumption is an error-ridden empirical measure of true consumption.

The pricing error associated with the true MRS, M_t^i , is by construction zero, but the pricing error associated with the erroneous MRS, M_t^c , for return R^j , is given by (dropping the time subscripts)

$$\text{Pricing error} \equiv PE = E [M^c R^j] - 1.$$

Throughout this paper, when we refer to pricing errors, we mean the pricing error generated for any asset by erroneously using the “pricing kernel,” M^c in place of the true pricing kernel, here M^i , since only the former are potentially nonzero.

How large are the pricing errors associated with using M^c in place of M^i ? To answer this question, first note that, for any asset return indexed by j , this pricing error can be rewritten in terms of log variables as

$$PE^j = E [\exp \{m^c + r^j\}] - 1,$$

which, given lognormality of C_t/C_{t-1} and returns, equals

$$PE^j = E [R^j] E [M^c] \exp \{ \text{Cov} (m^c, r^j) \} - 1.$$

Use the fact that the pricing error is identically zero under M^i to write

$$E [R^j] E [M^i] \exp \{ \text{Cov} (m^i, r^j) \} = 1,$$

implying that

$$PE^j = \frac{E [M^c]}{E [M^i]} \exp \{ \text{Cov} (m^c, r^j) - \text{Cov} (m^i, r^j) \} - 1 \quad (9)$$

$$= \frac{E [M^c]}{E [M^i]} \exp \{ -\gamma_c \text{Cov} (\Delta c, r^j) + -\gamma_i \text{Cov} (\Delta c^i, r^j) \} - 1. \quad (10)$$

This condition must hold for each asset, so that if the pricing errors for the risk-free rate and risky return associated with M^c were to be zero, we must have

$$\frac{E [M^c]}{E [M^i]} \exp \{ -\gamma_c \text{Cov} (\Delta c, r^s) + -\gamma \text{Cov} (\Delta c^i, r^s) \} = 1 \quad (11)$$

$$\frac{E [M^c]}{E [M^i]} \exp \{ -\gamma_c \text{Cov} (\Delta c, r^f) + -\gamma \text{Cov} (\Delta c^i, r^f) \} = 1. \quad (12)$$

It is now straight forward to solve for values of γ_c and δ_c that make equations (11) and (12) hold and therefore insure that the pricing errors $PE^f = PE^s = 0$. The resulting solutions are

$$\gamma_c = \gamma \left(\frac{\sigma_{is} - \sigma_{if}}{\sigma_{cs} - \sigma_{cf}} \right), \quad (13)$$

where $\sigma_{is} \equiv \text{Cov}(\Delta c^i, r^s)$, $\sigma_{cs} \equiv \text{Cov}(\Delta c, r^s)$ and σ_{if} and σ_{cf} are defined analogously for the risk-free return. The value of δ_c which satisfies (11) and (12) can be found by plugging (13) into either equation to find

$$\begin{aligned} \delta_c &= \delta \exp \left[\gamma_c \mu_c - \frac{\gamma_c^2 \sigma_c^2}{2} - \gamma \mu_i + \frac{\gamma^2 \sigma_i^2}{2} + \gamma_c \sigma_{cs} - \gamma \sigma_{is} \right] \\ &= \delta \exp \left[\gamma_c \mu_c - \frac{\gamma_c^2 \sigma_c^2}{2} - \gamma \mu_i + \frac{\gamma^2 \sigma_i^2}{2} + \gamma_c \sigma_{cf} - \gamma \sigma_{if} \right], \end{aligned} \quad (14)$$

where μ_c is the mean growth rate of aggregate consumption, and μ_i is the mean growth rate of the consumption of asset-holder i .⁶

Ruling out the knife-edge case in which $\sigma_{cs} = \sigma_{cf}$, equations (13) and (14) show that one can always find values of γ_c and δ_c that make the aggregate consumption-pricing errors associated with a broad stock return index and Treasury bill rate zero and indeed, these are the values an econometrician would find if the data were generated in equilibrium by (6) but one were to fit data to (8). This means that an erroneous pricing kernel based on aggregate consumption can always be found that unconditionally prices these two assets just as well as the true pricing kernel based on assetholder consumption. The estimates of γ_c and δ_c that result from fitting data to (8) will not correspond to any marginal individual's true risk-aversion or time discount factor. But a representative-agent pricing kernel based on per capita aggregate consumption can nevertheless be found that has the same unconditional asset-pricing implications as the true pricing kernel based on individual assetholder consumption.

At this point it should be clear that one would get identical results for *any* pricing kernel M_t^i that is jointly lognormally distributed with returns and aggregate consumption growth. It is not necessary that the pricing kernel take the form given in (6). Referring to (9) it is evident that the resulting solutions for δ_c and γ_c would be a function of the means, variances and covariances of Δc_t , r_t^s , r_t^f and m_t^i , whatever form the latter may take. If the true pricing

⁶Notice that, in equilibrium, γ_c and δ_c will take the same value regardless of the identity of the assetholder. This follows because any two households must in equilibrium agree on asset prices, so that the Euler equation holds for each individual household. Thus,

$$\gamma_c = \gamma_i \left(\frac{\sigma_{is} - \sigma_{if}}{\sigma_{cs} - \sigma_{cf}} \right) = \gamma_k \left(\frac{\sigma_{ks} - \sigma_{kf}}{\sigma_{cs} - \sigma_{cf}} \right)$$

for any two asset-holders i and k , as long as

$$E [M_t^i R_t^s] = E [M_t^k R_t^s]$$

and

$$E [M_t^i R_t^f] = E [M_t^k R_t^f].$$

kernel that sets Euler equation errors to zero is jointly lognormally distributed with aggregate consumption and returns, then values for the discount factor relative risk aversion can be always be found such that the standard model generates identical unconditional asset pricing implications for two asset returns, a risky and risk-free asset.

The solution for γ_c given above can be expressed in a more intuitively appealing way. Consider an orthogonal decomposition of aggregate consumption growth into a part that is correlated with asset-holder consumption and a part, ε_t^i , orthogonal to asset-holder consumption:

$$\Delta c_t = \beta \Delta c_t^i + \varepsilon_t^i, \quad (15)$$

where $\beta = \frac{\text{Cov}(\Delta c_t, \Delta c_t^i)}{\text{Var}(\Delta c_t^i)} \equiv \frac{\rho_{ci} \sigma_c}{\sigma_i}$. Here ρ_{ci} denotes the correlation between Δc_t and Δc_t^i . Using this decomposition, (13) can be re-written as

$$\gamma_c = \frac{\gamma}{\beta + \frac{\sigma_{\varepsilon^i s} - \sigma_{\varepsilon^i f}}{\sigma_{is} - \sigma_{if}}}, \quad (16)$$

where $\sigma_{\varepsilon^i s} \equiv \text{Cov}(\varepsilon_t^i, R_t^s)$ and $\sigma_{\varepsilon^i f} \equiv \text{Cov}(\varepsilon_t^i, R_t^f)$. Now consider assets that are uncorrelated with ε_t^i , the component of aggregate consumption that is uncorrelated with stockholder consumption. Many assets are likely to be included in this category, for example a broad stock market index or short-term interest rates, as those returns are unlikely to vary in a way that is orthogonal to assetholder consumption. Also included will be any risky asset that is on the log mean-variance efficient frontier.⁷ In this case we would have $\sigma_{\varepsilon^i s} = \sigma_{\varepsilon^i f} = 0$, and therefore

$$\gamma_c = \frac{\gamma}{\beta} = \gamma \frac{\sigma_i}{\rho_{ci} \sigma_c}. \quad (17)$$

Once we plug (17) into the formulas above for δ_c , we have solutions for δ_c and γ_c for which M_t^c perfectly satisfies the unconditional asset pricing restrictions of a risk-free return, and *any* log mean-variance efficient return. The formula tells us that the value of γ_c that would be obtained from fitting data to the erroneous kernel (8) will be higher the higher is assetholder risk aversion, the higher is assetholder consumption volatility relative to that of aggregate consumption, and the lower is the correlation between aggregate consumption growth and asset-holder consumption growth. Thus, limited participation and/or incomplete consumption insurance can in principal account for implausibly high estimated values of γ_c and δ_c obtained when fitting data to (8), but to do so, assetholder consumption must be

⁷This follows because the log return on any risky asset indexed by s can always be decomposed into a component that is correlated with the true log pricing kernel, m^i , and a component that is orthogonal to m^i , call it η^s . For any risky asset s , the covariance $\sigma_{\varepsilon^i s}$ will equal zero if and only if $\text{Cov}(\eta^s, \varepsilon^i) = 0$. Naturally the latter will hold if the variance of η^j is zero, which in turn will occur if the correlation between m , and r^j is -1, that is R_t^s is on the log mean-variance efficient frontier.

more volatile than aggregate consumption and/or very weakly correlated with it. Notice, however, that even if assetholder consumption behaves very differently from per capita aggregate consumption, this is not enough to explain the large unconditional Euler equation errors that arise from fitting data to (8). The only consequence of using aggregate per capita consumption in this setting is a bias in the estimate parameters γ_c and δ_c ; there is no consequence for the Euler equation errors, which remain zero.

3.2 Leading Asset Pricing Models

In this section we consider the Euler equation errors generated by specific models. Can the large unconditional Euler equation errors generated by fitting data to (1) be explained by leading asset pricing models? Does discarding the standard pricing kernel in favor of the true kernel implied by these models allow an econometrician to better model asset pricing data? All of the leading models we consider are consumption-based asset pricing models specifically designed to resolve puzzles associated with the standard model. In addition, all of the leading models develop endogenous predictions for a stock market return (sometimes modeled as the return to an aggregate wealth portfolio) and a risk-free rate, and none imply that the pricing kernel is unconditionally jointly lognormally distributed with aggregate consumption growth and returns.⁸ We now turn our attention to investigating each model’s implications for the unconditional pricing errors. We do so by fitting artificial data generated in equilibrium by each model to (1).

3.2.1 Misspecified Preferences

Suppose (1) is fitted to data generated by a representative agent model with preferences that differ from power utility. Can leading asset pricing models with different preferences explain the large empirical pricing errors found in Section 2? We consider three prominent representative-agent models: the external habit-persistence models of Campbell and Cochrane (1999) (CC) and Menzly, Santos, and Veronesi (2004) (MSV), and the long-run risks model of Bansal and Yaron (2004).⁹ All three of these models display a striking ability to match a range of asset pricing phenomena, including a high equity premium, low and stable risk-free rate, long-horizon predictability of excess stock returns, and countercyclical variation in the Sharpe ratio (where the Sharpe ratio is defined as the conditional mean

⁸Joint lognormality of consumption growth, the risky asset return and the risk-free return can be statistically rejected in simulated data of all of the models discussed in this section.

⁹The habit models generate time-varying and thus state-dependent risk-aversion. Melino and Yang (2003) and Gordon and ST-Amour (2004) consider more general models of state dependent preferences to study asset pricing phenomenon.

of the excess stock market return divided by its conditional standard deviation). In what follows, we describe only the main features of each model, and refer the reader to the original article for details. Except where noted, our simulations use the baseline parameter values of each paper.

The utility function in the CC and MSV models take the form

$$U = E \left\{ \sum_{t=0}^{\infty} \delta^t \frac{(C_t - X_t)^{1-\gamma} - 1}{1-\gamma} \right\}, \quad \gamma > 0 \quad (18)$$

where C_t is individual consumption and X_t is habit level which they assume to be a function of aggregate consumption, and δ is the subjective discount factor. In equilibrium, identical agents choose the same level of consumption, so C_t is equal to aggregate consumption. The key innovation in each of these models concerns the specification of the habit process X_t , which in both cases evolves according to distinct heteroskedastic processes. (The Appendix provides a more detailed description of the models in this section.) The stochastic discount factors in both models take the form

$$M_{t+1} = \delta \left(\frac{C_{t+1} - X_{t+1}}{C_t - X_t} \right)^{-\gamma}$$

but differ in their specification of X_t . We denote as M_{t+1}^{CC} the specification of M_{t+1} corresponding to the Campbell-Cochrane model of X_t , and as M_{t+1}^{MSV} the specification of M_{t+1} corresponding to the MSV model of X_t .

CC and MSV assume that the log difference in consumption, $\Delta c_t \equiv \log(C_t/C_{t-1})$, follows an i.i.d. process:

$$\Delta c_t = \mu + \sigma v_t,$$

where v_t is a normally distributed i.i.d. shock. Both models derive equilibrium returns for a risk-free asset and a risky equity claim that pays aggregate consumption as its dividend. MSV also extend the Campbell and Cochrane model by considering the equilibrium pricing of multiple risky securities, but for the moment we focus on the model's implications for the stock return, R_{t+1}^s , and risk-free rate, R_{t+1}^f . Campbell and Cochrane set $\gamma = 2$ and $\delta = 0.89$ at an annual rate. Menzly, Santos and Veronesi choose $\gamma = 1$ and $\delta = 0.96$. Notice that the curvature parameter γ , is no longer equal to relative risk-aversion.

Bansal and Yaron (2004) consider a representative-agent who maximizes utility given by recursive preferences of Epstein and Zin (1989, 1991) and Weil (1989). The stochastic discount factor under Epstein-Zin-Weil utility used in BY takes the form

$$M_{t+1}^{BY} = \left(\delta \left(\frac{C_{t+1}}{C_t} \right)^{-\frac{1}{\psi}} \right)^{\alpha} R_{w,t+1}^{\alpha-1}, \quad (19)$$

where $R_{w,t+1}$ is the simple gross return on the aggregate wealth portfolio, which pays a dividend equal to aggregate consumption, C_t , $\alpha \equiv (1 - \gamma) / (1 - 1/\psi)$, ψ is the intertemporal elasticity of substitution in consumption (IES), γ is the coefficient of relative risk aversion, and δ is the subjective discount factor. Bansal and Yaron assume that both aggregate consumption growth and aggregate dividend growth have a small predictable component that is highly persistent. They also incorporate stochastic volatility into the exogenous processes for consumption and dividends to capture evidence of time-varying risk-premia. Taken together, the dynamics of consumption growth and stock market dividend growth, Δd_t , take the form

$$\Delta c_{t+1} = \mu + x_t + \sigma_t \eta_{t+1} \quad (20)$$

$$\Delta d_{t+1} = \mu_d + \phi x_t + \rho_d \sigma_t u_{t+1}, \quad (21)$$

$$x_{t+1} = \rho x_t + \rho_c \sigma_t e_{t+1}$$

$$\sigma_{t+1}^2 = \sigma^2 + \nu_1 (\sigma_t^2 - \sigma^2) + \sigma_w w_{t+1},$$

where σ_{t+1}^2 represents the time-varying stochastic volatility, σ^2 is its unconditional mean, and μ , μ_d , ϕ , ρ_d , ρ , ρ_c , ν_1 and σ_w are parameters, calibrated as in BY. Here, the stock market asset is the dividend claim, given by (21), rather than a claim to aggregate consumption, given by (20). BY calibrate the model so that x_t is very persistent, with a small unconditional variance. Thus, x_t captures long-run risks, since a small but persistent component in the aggregate endowment can lead to large fluctuations in the present discounted value of future dividends. Their favored specification sets $\delta = 0.998$, $\gamma = 10$ and $\psi = 1.5$.¹⁰

For each model above, a solution is obtained and a large time series of artificial data (20,000 observations) are generated. We use these data to quantify the magnitude of unconditional pricing errors for the equilibrium risk-free rate and stock market return that would arise if an econometrician fit $M_t^c \equiv \delta_c (C_t/C_{t-1})^{-\gamma_c}$ to data generated by each of the models described above. The parameters δ_c and γ_c are chosen to minimize an equally-weighted sum of squared pricing errors,

$$\min_{\delta_c, \gamma_c} E \left[\delta_c (C_t/C_{t-1})^{-\gamma_c} R_t^s - 1 \right]^2 + E \left[\delta_c (C_t/C_{t-1})^{-\gamma_c} R_t^f - 1 \right]^2,$$

where the data on aggregate consumption, C_t , the stock return, R_t^s , and the risk-free rate, R_t^f are model-generated simulated data of the equilibrium outcomes of each model. Table 4 reports the results.

In each case, we find the pricing errors that arise from erroneously using the standard pricing kernel based on power utility are numerically zero, just as they are when the true

¹⁰The results below are based on the first-order approximate analytical solutions given in BY. The simulation results are close to those based on the numerical solutions reported in Bansal and Yaron (2004).

pricing kernel is used. Values of δ_c and γ_c can in each case be found that allow the standard consumption-based model to unconditionally price assets just as well as the true pricing kernel. In the CC model, the values of δ_c and γ_c that set these pricing errors to zero are 1.28 and 57.48 respectively. In the MSV model, the corresponding values are 1.71 and 30.64, respectively. Thus, the habit models *can* explain what many would consider the implausible estimates of time preference and risk aversion obtained when freely fitting aggregate data to (1). Recall that the true preference parameters are $\gamma = 2$ and $\delta = 0.89$ in CC and $\gamma = 1$ and $\delta = 0.96$ in MSV. But it is in those parameters that all of the distortion from erroneously using M_t^c arises. No distortion appears in the Euler equation errors themselves. In terms of obtaining the correct unconditional asset pricing implications, these models do not appear to be very important: an econometrician could model the asset pricing data just as effectively using a standard pricing kernel based on power preferences. It follows that the models cannot explain the large unconditional Euler equation errors that arise when fitting historical data to (1).

The findings are similar for the Bansal-Yaron long-run risks model. Here we follow BY and simulate the model at monthly frequency, aggregating to annual frequency to report the model's asset pricing implications. Thus, the monthly consumption data are time-aggregated to annual consumption, and monthly returns are compounded to annual returns.¹¹ We find that δ is close to the true value, but γ is estimated to be about five times as high as true risk aversion. Thus, as for the habit models, this framework explains why an econometrician obtains high estimates of risk aversion when fitting data to the standard consumption-based model. But, also like the habit models, if an econometrician fit M_t^c to data generated by M_t^{BY} , the resulting Euler equation errors would be effectively zero, in contrast to what is found using historical data.¹²

3.2.2 Misspecified Consumption

Suppose (1) is fitted to data generated by a non-representative agent model in which asset prices are determined not by per capita aggregate consumption but rather by the consumption of stockholders. Can asset pricing models in which stockholder consumption belongs in the pricing kernel explain the large empirical pricing errors computed from aggregate con-

¹¹The resulting Euler equation errors are unchanged if they are computed for quarterly time-aggregate consumption and quarterly returns rather than annual time-aggregated consumption and annual returns.

¹²For models based on recursive preferences, Kocherlakota (1990) shows that there is an observational equivalence to the standard model with power utility preferences, if the aggregate endowment growth is i.i.d. However, the endowment growth process in the BY model is not i.i.d., but instead serially correlated with stochastic volatility. Moreover, the annual consumption data are time-aggregated, which further distorts the time-series properties from those of the monthly endowment process.

sumption data, as in Section 2? We consider the limited participation model of Guvenen (2003) (GUV). Like the representative-agent models considered above, this Guvenen model has demonstrated remarkable success in explaining many of the empirical puzzles associated the standard representative-agent consumption-based model. It can account for a high equity premium and low and stable risk-free rate, predictable stock market returns, and countercyclical Sharpe ratio.

The Guvenen model has two types of consumers, stockholders and nonstockholders. The latter are exogenously prevented from participating in the stock market. The discount factor in this model is given by

$$M_{t+1}^{GUV} = \delta \left(\frac{C_{t+1}^i}{C_t^i} \right)^{-\gamma},$$

where δ is the subjective discount factor of the stockholder, γ is the stockholder's relative risk aversion, and C_t^i is stockholder consumption, which by assumption cannot be the same as aggregate per capita consumption. In other respects, the model is a standard one-sector real business cycle model with adjustment costs in capital. Both stockholders and nonstockholders receive labor income with wages determined competitively by the marginal product of labor, and firms choose output by maximizing the present discounted value of expected future profits. Both agents have access to a riskless bond.

We follow the same procedure discussed above to quantify pricing errors in this model. We simulate a large time series of artificial data (20,000 observations) for the equilibrium values of the variables in this model, and use these data to quantify the magnitude of unconditional pricing errors that an econometrician would find from fitting (1) to data generated by M_{t+1}^{GUV} .

In Guvenen's baseline model, stockholders have risk aversion $\gamma = 2$ and subjective discount factor $\delta = 0.99$. Table 5, panel A shows that stockholder consumption growth is about two and a quarter times as volatile as aggregate consumption growth in the baseline model, and perfectly correlated with it. The model implies that stockholder consumption is over four times as volatile as nonstockholder consumption growth, but the two are also highly correlated, with correlation 0.99. This is not surprising since both types of consumers participate in the same labor market and bond markets; the agents differ only in their ability to hold equities, and in their risk-aversion (nonstockholders have higher risk-aversion). As a consequence, stockholder's MRS, M_t^{GUV} , is highly correlated with the aggregate consumption pricing kernel $M_t^c \equiv \delta_c (C_t/C_{t-1})^{-\gamma_c}$, for a variety of values of δ_c and γ_c . Panel B of Table 5 shows this correlation for two such cases, first with $\delta_c = \delta = 0.99$ and $\gamma_c = \gamma = 2$, and second with δ_c and γ_c set to the values that minimize the equally-weighted sum of squared pricing errors for the stock return and riskless bond. These latter values are $\delta_c = 0.99$ and $\gamma_c = 4.49$. In both cases, the correlation between M_t^{GUV} and M_t^c is extremely high, 0.99.

Panel C of Table 5 shows the pricing errors in Guvenen's model that would arise if

aggregate consumption were erroneously used in the pricing kernel in place of stockholder consumption. For comparison, the table also reports the pricing errors using the true kernel M_t^{GUV} based on stockholder consumption, which are quite small (0.02% on an annual basis) but not exactly zero due to the rarely-binding borrowing constraints that apply to both stockholders and nonstockholders. When $\delta_c = \delta = 0.99$ and $\gamma_c = \gamma = 2$, the pricing errors using aggregate consumption are larger than that using stockholder consumption, equal to about 0.4% at an annual rate for the stock return and -0.34% for the risk-free rate. Although the signs of the errors match those in the data, they are much smaller in magnitude. When δ_c and γ_c are chosen to minimize the sum of squared pricing errors for these two asset returns, as in empirical practice, the pricing errors are, to numerical accuracy, zero for the stock return and a mere 0.01% at an annual rate for the risk-free return. Thus, the model delivers a pricing kernel using aggregate consumption that is virtually identical in its unconditional pricing implications for these two moments to the true pricing kernel using stockholder consumption. While the model produces small pricing errors using the true pricing kernel based on stockholder consumption, like the models above, it counterfactually produces small pricing errors using the wrong pricing kernel M_t^c based on per capita aggregate consumption. We are led to the conclusion that Euler equation errors of the standard consumption-based model cannot be attributable to merely using the wrong consumption measure. Moreover, in this model, stockholder consumption is irrelevant for getting the average price of these assets correct and does not explain the high estimated values of δ and γ obtained when fitting data to the standard consumption-based model. Instead, with a modest adjustment to the risk-aversion parameter, there is an observational equivalence between the standard consumption-based pricing kernel and the true pricing kernel based on stockholder consumption. An econometrician who fits M_t^c to data generated by M_t^{GUV} would fail to reject—either on statistical or economic grounds—the standard consumption-based model.

3.2.3 Misspecified Preferences and Misspecified Consumption

One possible reaction to the results above, at least for the representative agent models, is that we should take the representative-agent nature of the models less literally and assume that they apply only to a representative stockholder, rather than to a representative household of all consumers. Would the results for these models be better reconciled with the data if we accounted for limited participation? Not necessarily. As an illustration, we consider a simple limited-participation version of the MSV model and show that the results are essentially unchanged from the representative-agent setup.

Since the MSV model is a representative-agent model, we modify it in order to study the role of limited participation. Assume that asset prices are determined by the framework

above, where a valid stochastic discount factor is a function of any stockholder's consumption C_t^i and stockholder's habit X_t^i . The process for stockholder consumption is the same as in MSV, described above, but now with i subscripts:

$$\Delta C_t^i = \mu_i + \sigma_i v_t^i,$$

where v_t^i is a normally distributed i.i.d. shock. Aggregate consumption is assumed to follow a separate process given by

$$\Delta C_t = \mu_c + \sigma_c v_t^c,$$

with v_t^c a normally distributed i.i.d. shock. We analyze the results over a range of cases for the correlation between v_t^i and v_t^c , and their relative volatilities σ_i/σ_c .

Asset prices are determined by the stochastic discount factor of assetholder defined as $M_{t+1}^{MSVi} = \delta \left(\frac{C_{t+1}^i - X_{t+1}^i}{C_t^i - X_t^i} \right)^{-\gamma}$, where X_{t+1}^i is the habit modeled as in MSV. Next, we compute two types of unconditional pricing errors. First, we compute the pricing errors generated from erroneously using aggregate consumption in the pricing kernel in place of assetholder consumption. That is, we compute the pricing errors that arise from using $M_{t+1}^{ch} \equiv \delta^c \left(\frac{C_{t+1} - X_{t+1}}{C_t - X_t} \right)^{-\gamma^c}$ in place of M_{t+1}^{MSVi} to price assets, where X_t is computed from the MSV habit specification using aggregate consumption, and where δ^c and γ^c are chosen to minimize the pricing errors. Second, we compute the unconditional pricing errors from *both* erroneously using aggregate consumption in the pricing kernel, and from erroneously using power preferences. That is, we compute the pricing errors that arise from using $M_t^c = \delta_c (C_t/C_{t-1})^{-\gamma_c}$ to price assets when the true model is M_{t+1}^{MSVi} . As above, the parameters δ_c and γ_c are chosen to minimize an equally-weighted sum of squared pricing errors.

The results are presented in Table 6 for the exercise using M_{t+1}^{ch} and Table 7 for the exercise using M_t^c . We report the pricing errors associated with erroneously using M_{t+1}^{ch} and M_{t+1}^c in place of M_{t+1}^i , for a range of values for the relative volatility between asset-holder and aggregate consumption growth and for their correlations. The standard deviation of asset-holder consumption growth is allowed to range from one times to ten times as volatile as that of aggregate consumption growth. The correlation is allowed to range from -1.0 to 1.0. The pricing errors are reported in the bottom panels for the return to the aggregate wealth portfolio, denoted R_{t+1}^s and for the risk-free rate, R_t^f . We also report the values of δ_c and γ_c that set these errors as close to zero as possible, for each case.

Table 6 shows that the pricing errors that arise from using M_{t+1}^{ch} to price assets are always zero. This shows that even if preferences are not of the standard iso-elastic form, using the wrong consumption measure in the pricing kernel does not necessarily generate large pricing errors, even if assetholder consumption is five times as volatile as aggregate consumption,

or very weakly correlated with it. Instead, a stochastic discount factor using aggregate consumption can be found that is virtually identical in its pricing implications to the true pricing kernel in the same class of preferences. Notice that the parameters δ_c and γ_c can deviate substantially from the true values. This reinforces the findings from above, which suggest that merely making assetholder consumption growth behave very differently from aggregate consumption growth is not enough to explain large pricing errors generated by the standard consumption-based model. Table 6 shows that aggregate consumption growth can be perfectly negatively correlated with assetholder consumption growth and five times as volatile, yet the pricing errors that arise from using C_t in place of C_t^i are still zero. Instead, all of the adjustment is loaded into the coefficients, δ^c and γ^c .

Table 7 shows that the same qualitative result holds if one uses M_{t+1}^c in place of the true pricing kernel M_{t+1}^i . Here we use *both* the wrong consumption measure and the wrong preferences. There are a few cases for which the numerical procedure does not converge, denoted “NA” in the table. Except for these cases, we can as before find values of δ^c and γ^c such that M_{t+1}^c explains the Euler equations just as well as M_{t+1}^i . The values for δ^c and γ^c are more distorted from their true values than is the case in Table 6 where we have merely substituted the wrong consumption measure into the class of habit preferences, but the pricing errors of the standard consumption-based model are still zero.¹³ These findings reinforce the conclusion that changing the pricing kernel does not necessarily change the pricing implications.

3.2.4 Larger Cross-Sections of Returns

So far, we have only considered the pricing implications of two asset returns, a risk-free rate and a risky asset return. Of the models above, three of them, namely the Campbell-Cochrane, Bansal-Yaron, and Guvenen models, generate implications only for these two asset returns,¹⁴ returns that are the focal point of those studies. By contrast, the model of Menzly, Santos and Veronesi generates implications for many risky securities, each distinguished by a distinct dividend process whose dynamics are characterized by fluctuations in the share it represents in aggregate consumption. Of course, the model also generates implications for the risk-free rate and aggregate wealth portfolio return, the implications of which were analyzed above. We now exploit the structure of the MSV model to study its implications

¹³Notice that variation in σ_i/σ_c has little affect on the estimated value of the risk-aversion parameter γ_c . This happens because we adjust α at the same time as we adjust σ_i/σ_c so that the mean excess return remains roughly what it is in MSV. Since the volatility of aggregate consumption is kept the same and α is adjusted to keep the returns of the same magnitude, γ_c doesn't change much.

¹⁴In principal, one could study the implications of these models for *scaled* returns, which are the raw returns multiplied by some known conditioning variable given in equilibrium by the model.

for a larger cross-section of equilibrium returns, analogous to the empirical exercise in Panel B of Table 1 using historical data on 8 asset returns.

MSV model the share of aggregate consumption that each asset produces,

$$s_t^j = \frac{D_t^j}{C_t} \quad \text{for } j = 1, \dots, n,$$

where n represents the total number of risky financial assets paying a dividend D . The Appendix gives a more detailed description of the stochastic process MSV assume for the shares and our calibration of the model. Cross-sectional variation in unconditional mean returns across risky securities in this model is governed by cross-sectional variation in the covariance between shares and aggregate consumption growth. In analogy to the empirical exercise (Panel B of Table 1), we create six artificial risky securities plus the aggregate wealth portfolio return and the risk-free for a total of 8 asset returns. The parameters of the share processes are chosen to generate a spread in expected returns across assets.

We simulate 20,000 periods and compute the pricing errors that would arise in equilibrium if $M_{t+1}^c = \delta_c \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma_c}$ were used to price assets in place of the true kernel M_t^{MSV} . Since we now have a total of 8 asset returns, we choose the parameters γ_c and δ_c to minimize the quadratic form $g_T(\gamma, \delta) \equiv \mathbf{w}'_T(\gamma, \delta) \mathbf{W} \mathbf{w}_T(\gamma, \delta)$, where $\mathbf{w}_T(\gamma, \delta)$ is the (8×1) vector of sample average pricing errors using the model-simulated data (i.e., $w_{jT}(\gamma, \delta) = \frac{1}{T} \sum_{t=1}^T \delta \left(\frac{C_t}{C_{t-1}} \right)^{-\gamma} R_t^j - 1$ for $j = 1, \dots, 8$) and \mathbf{W} is the 8×8 identity matrix. We maintain the same limited participation set-up described above for the two-asset results reported in Tables 5 and 6 and report the pricing errors associated with erroneously using M_{t+1}^c in place of the true kernel M_{t+1}^{MSVi} for a range of values for the relative volatility between asset-holder and aggregate consumption growth and for their correlations. The results are reported in Table 8.

The root-mean-squared pricing error, as a percentage of the cross-sectional mean return, is reported in the bottom panel. We also report the values of δ_c and γ_c that minimize the quadratic criterion function, for each case. The results here are qualitatively the same as those in Table 7 for two-asset returns: the estimation produces wildly inaccurate values for δ and γ , but the standard model pricing kernel M_{t+1}^c can explain returns virtually as well as the true pricing kernel M_{t+1}^{MSVi} . The root-mean-squared-error that arises from erroneously using M_{t+1}^c to price assets is tiny as a fraction of the cross-sectional mean return, the highest being about 4%. These numbers should be contrasted with the 35% figure obtained for a cross-section of 8 asset returns in U.S. data (Table 1, Panel B). Moreover, the numbers in Table 8 actually overstate the true pricing errors. This is because there are two sources of error that result in nonzero pricing errors even using the true pricing kernel M_{t+1}^{MSVi} . The first is the discrete-time approximation to the continuous-time model of MSV. We eliminate much, but

not all, of this error by shrinking the time-interval over which we simulate the model and reporting annualized values in the table. The second source of error is the approximation in (25). Taken together, these errors mean that the true kernel generates pricing errors that are often of the same order of magnitude as those reported in Table 8. These results reinforce the conclusions from the two-asset case and demonstrate that the unconditional pricing implications of the standard consumption-based model can easily be virtually identical to those using the true pricing kernel, even when the number of assets exceeds the number of free parameters and when both the wrong preferences and the wrong consumption measure are employed to price assets.

3.2.5 Time Aggregated Consumption

With the exception of the Bansal-Yaron model, the results presented above use artificial data generated from the original models in which the decision interval of agents is equal to the data sampling interval. One possible concern about the results so far is that the decision interval of households may be shorter than the data sampling interval, leading to time-aggregated consumption observations. Since time-aggregation is likely to affect the correlation between returns and consumption growth, it is important to assess its influence on pricing errors. We have repeated the same exercise for all the models above using time-aggregated consumption data, assuming that agents' decision intervals are shorter than the data sampling interval. For example, we assume that agents make decisions quarterly but that the data sampling interval is annual. In this case, the annual observation on consumption is the time-aggregated value of consumption over the four quarters in a year, where the annual return is the compounded equilibrium quarterly return. For all models the essential results for the Euler equation errors remain unchanged: values of δ_c and γ_c can always be found such that the unconditional pricing errors associated with using M_t^c to price assets are numerically zero, even when using time-averaged data. We demonstrate this in Table 9, for the MSV model with limited participation. (To conserve space, we report only the results for this model, since the conclusion is unchanged for the other models.) Since time-averaging changes both the serial dependence of the consumption data and its unconditional correlation with returns, this suggests that the exact time-series properties of consumption growth is often not crucial for explaining the empirical facts presented here.

To summarize, if any of the leading models we consider above were true, we would not observe the pricing errors we observe when fitting data to the standard consumption-based asset pricing model. Instead, Euler equation errors would be zero and the entire consequence of using the wrong pricing kernel would be incorrect estimates of δ and γ . If data were generated from any of the models M_t^{CC} , M_t^{MSV} , M_t^{BY} , and M_t^{GUV} , an econometrician would

find that they have no more explanatory power than the standard representative-agent, consumption-based asset pricing model in tests of unconditional asset pricing restrictions; put another way, an econometrician would never statistically reject the standard consumption-based model. To think about one way such large unconditional pricing errors might be explained in frictionless models, we now dig a bit deeper into the potential roles of limited participation/incomplete markets. As stated, the common defining feature of both types of models is the presumption that it is not aggregate per capita consumption that is required to explain asset returns, but rather the consumption of some subset of the aggregate who are assetholders. We saw above that when all variables are jointly lognormally distributed, there is no scope for expanding the unconditional asset pricing implications of the standard model with limited participation or incomplete markets. The next section considers what happens when these variables are allowed to depart from lognormality.

3.3 Perturbations from Normality: Limited Participation/Incomplete Markets

How do the unconditional pricing implications of models with limited participation/incomplete markets change when variables are not jointly lognormal? We approach this question by considering distributions that are perturbed from the normal thereby allowing departures from normality. In particular, we restrict our investigation to first-order Hermite expansions around the normal. We do so for three reasons. First, as we explain below, such a restriction makes the analysis tractable given the number of unknown parameters that must be calibrated. Second, the first-order expansion can accommodate significant departures from normality, if not arbitrary departures. Third, we argue that at least a large number of commonly parameterized economic models built from the primitives of tastes, technology and fundamental shocks are likely to generate endogenous variables that fall within this class of distributions. One caveat is that the distributions we consider cannot accommodate conditional heteroskedasticity or other forms of conditional temporal dependence. Allowing for such dependence would require the calibration of an infeasible number of Hermite parameters about which we have no information. Still, we will show that it is not necessary to entertain such generalizations in order to find examples of models capable of matching the empirical regularities we focus on here. We begin this section by considering the Euler equation errors associated with a stock market return and a risk-free rate and later move on to consider a larger cross-section of asset returns.

Let $y_t = (\Delta c_t, \Delta c_t^i, \Delta d_t)' \equiv (y_{1,t}, y_{2,t}, y_{3,t})'$, a vector containing aggregate consumption growth, Δc_t , individual asset-holder consumption growth, Δc_t^i , and dividend growth on the aggregate stock market return, Δd_t . We will consider simple asset pricing models in which

these variables are i.i.d., but not necessarily jointly lognormally distributed.

Let the joint density of y_t be denoted $h(y)$. How can $h(y)$ be chosen without imposing potentially wildly erroneous shape? We use a Hermite polynomial expansion around the normal, which is a polynomial in y times the standard Gaussian density. The primary advantage of the Hermite expansion is that the leading term is Gaussian, and higher-order terms accommodate deviations from normality. Gallant and Tauchen (1989) show that such an expansion can be put in tractable form by specifying the density as

$$h(y) = \frac{a(y)^2 f(y)}{\int \int \int a(u)^2 f(u) du_1 du_2 du_3}.$$

Here, $f(y)$ is the multivariate Gaussian density with variance-covariance matrix Ω and mean $\mu = (\mu_1, \mu_2, \mu_3)'$, and $a(y)$ is the sum of polynomial basis functions of the variables in y ; it is squared to insure positivity of the density and divided by the integral over \mathbb{R}^3 to insure the density integrates to unity.

In our calibrated examples, we set $a(y)^2 = (a_0 + a_1 y_{1,t} + a_2 y_{2,t} + a_3 y_{3,t})^2$, a low-order expansion but one that can nonetheless accommodate quite significant departures from normality. We can investigate results for a large number of possible joint distributions by simply varying the parameters a_0, \dots, a_3 . When $a_0 = 1$ and $a_1 = a_2 = a_3 = 0$, $h(y)$ collapses to the Gaussian joint distribution, $f(y)$. It is important to keep the degree of the Hermite expansion manageable since, lacking a sufficiently long times series on asset-holder consumption, we cannot feasibly estimate the parameters of $f(y)$ and $a(y)$.¹⁵ Hence we also restrict the variables to be i.i.d., since accounting for variation in conditional distributions would increase the required number of hermite parameters by a factor of ten. As it is, the current set-up already leaves a number of parameters to be calibrated, including those in Ω , μ , and $a(y)$. Given that we have little direct information about many of these parameters, we will look at a range, leaving a large number of parameter combinations to be calibrated. We discuss this calibration below.

As before, we denote the risky return with an s superscript and the risk-free rate with an f superscript. In equilibrium we must have

$$E_t [M_{t+1}^i R_{t+1}^j] = 1, \quad j = s, f,$$

where $M_{t+1}^i \equiv \delta (C_{t+1}^i / C_t^i)^{-\gamma}$ is the true pricing kernel based on individual assetholder consumption. For the stock return, R_{t+1}^s , this expression may be written as

$$E_t \left[M_{t+1}^i \left(\frac{P_{t+1}}{D_{t+1}} + 1 \right) \frac{D_{t+1}}{D_t} \right] = \frac{P_t}{D_t}, \quad (22)$$

¹⁵Gallant and Tauchen (1989) estimate a number of Hermite polynomial densities for aggregate consumption growth, stock returns and a Treasury-bill rate, for which a long time-series of data are available.

where P_t is the end-of-period stock price at time t and we have used the definition of gross returns, $R_{t+1}^s \equiv (P_{t+1} + D_{t+1})/P_t$. For the risk-free rate, an analogous equation holds using the definition $R_{t+1}^f \equiv (E_t [M_{t+1}^i])^{-1}$, but notice that since all variables are i.i.d., conditional expectations are just the same as unconditional expectations. Thus, we can use the unconditional joint distribution $h(y)$ to compute the expectation in (22) and the equilibrium risk-free rate. Because all variables are i.i.d., the equilibrium price-dividend ratio is a constant, P/D , implying that the stock return's distribution is known and is the same as the distribution of dividend growth. We solve for the price-dividend ratio P/D that satisfies (22), from which one can compute the equilibrium stock return. The solution for the price-dividend ratio satisfies the equation

$$\frac{P/D}{P/D + 1} = \int \int \delta^i \exp(-\gamma^i y_2) \exp(y_3) h(y_2, y_3) dy_2 dy_3,$$

where y_2 and y_3 correspond to asset-holder consumption growth and aggregate dividend growth. Given a distribution $h(y)$ and the equilibrium value for P/D , it is straightforward to compute the pricing errors for the risky asset return and risk-free return associated with erroneously using $M_t^c \equiv \delta_c (C_t/C_{t-1})^{-\gamma_c}$ to price assets:

$$PE_k = E [M_{t+1}^c R_{t+1}^k] - 1 \quad k = s, f. \quad (23)$$

We solve numerically for the values of δ_c and γ_c that minimize the pricing errors (23) associated with the stock return R_t^s and risk-free rate R_t^f .

For our numerical computations, where possible, the parameters of the leading normal density $f(y)$ are calibrated to match data on aggregate consumption growth and dividend growth for the CRSP value-weighted stock market index, on an annual basis. We take the mean of Δc to be 2% annually and the mean of Δd to be 4% annually from annual post-war data used in Lettau and Ludvigson (2004). From the same annual data, the standard deviation of aggregate consumption growth is $\sigma_c = 1.14\%$ and the standard deviation of dividend growth is $\sigma_d = 12.2\%$. The covariance between Δc and Δd , denoted σ_{cd} , is notoriously hard to measure accurately. It is estimated to be small and negative, equal to -0.000177 in the annual data used by Lettau and Ludvigson (2004), but others have estimated a weak positive correlation (e.g., Campbell (2003)). We therefore consider both small negative values for this covariance (equal to the point estimate from Lettau and Ludvigson (2004)), and small positive values of the same order of magnitude, e.g., 0.000177. Finally, the parameters for asset-holder consumption and asset holder preferences must be calibrated since we lack a sufficiently long time series on individual consumption to measure these parameters with any reasonable degree of accuracy. We therefore consider a range for γ , δ , σ_i/σ_c , μ_i/μ_c , ρ_{ci} , and ρ_{id} , where ρ_{id} is the correlation between asset-holder consumption growth and dividend

growth. Because our calibration corresponds to an annual frequency, the Euler equation errors are comparable to the annualized errors from U.S. data reported in Table 1.

We begin with an example of a joint distribution that can roughly replicate the large Euler equation errors that arise from fitting the data to (1). We stress that this is only one example, but within the class of distributions we investigate here, it seems to be representative of what is required. Clearly distributions outside this class could provide other examples of models capable of explaining the unconditional Euler equation facts documented in Section 2. At the same time, while this example is meant to be illustrative, we stress that it does not necessarily constitute a plausible resolution to empirical regularities presented in Section 2. Such a resolution would require much more evidence on the joint distribution of aggregate consumption, assetholder consumption, and asset returns than what is currently available.

The marginal distributions for Δc , Δc^i , and Δd for this example are presented in Figure 5. The parameters in the leading normal are set as follows: $\sigma_i/\sigma_c = 4$, $\mu_i/\mu_c = 1.5$, $\rho_{ci} = 0.1$, and $\rho_{id} = 0.9$. Assetholder risk aversion is set to a moderate value of $\gamma = 5$ and the time discount factor is set to $\delta = 0.99$. The other parameters were calibrated as described above, to match aggregate data on consumption and dividends. The Hermite parameters a_0, \dots, a_3 , are set to obtain the density shapes displayed in Figure 5.

For this particular joint distribution model, the pricing error that arises from erroneously using $M_t^c \equiv \delta_c (C_t/C_{t-1})^{-\gamma_c}$ to price assets is 2.81% for the stock return, and -2.99% for the risk-free rate. Notice that these values are close to those in the historical data for the CRSP stock return and 3-month Treasury bill rate (Table 1). The average stock return in this example is about 11%, and the average risk-free rate 4% annually. These are both a bit higher than in the historical data, but the pricing errors as a fraction of the average returns are reasonably close to the data. The resulting standard deviation of Δd is a bit higher and its mean a bit lower than the corresponding figures for the CRSP-VW return, but the example is nonetheless instructive.

What features distinguish this example? First, notice from Figure 5 that assetholder consumption growth and the risky return are highly correlated with one another, but neither is highly correlated with aggregate consumption growth. Second, both assetholder consumption growth and dividend growth are much more volatile than aggregate consumption growth, with the former six times as volatile as that of aggregate consumption growth. Third, the density of aggregate consumption is almost identical to the leading normal with its mean and variance calibrated from the data. By contrast, stockholder consumption and dividend growth have distributions that differ significantly from normality, with both displaying bimodal densities. Assetholder consumption and dividend growth have about equal mass points at steeply negative and positive growth rates not present in the density of aggregate consumption. With probability 0.25, assetholder consumption can decline by 5%,

while such a steep decline receives no weight in the aggregate consumption density. Similarly, with about 0.2 probability, assetholder consumption can grow as fast as 10% while dividend growth on the risky asset can grow about 25%, again zero-probability events for aggregate consumption growth. Simulations from such a distribution would deliver periods in which the joint behavior of M_t^c and returns would be quite different from the joint behavior of M_t^i and returns. This model produces observations that explain the large pricing errors that arise when fitting aggregate data to (1).

Within the class of models we consider, how common is this example? To address this question, we evaluated pricing errors obtained from over 20,000 parameter combinations on a wide grid for the hermite parameters a_0 through a_3 . Since it is infeasible to report the output from tens of thousands of distributional assumptions, we report a limited number of parameter combinations in Table 10 that illustrate a range of different shapes of densities, skewness, and kurtosis statistics. Two restrictions place significant limitations on the number of valid parameter combinations that can be considered. First, the parameters ρ_{ci} and ρ_{id} cannot be given arbitrary values or Ω will not be positive semi-definite. In particular, if the correlation between dividend growth and assetholder consumption growth is high while the correlation between aggregate consumption growth and dividends is low, the correlation between aggregate consumption growth and assetholder consumption growth cannot be too high. Since the data suggest a weak correlation between aggregate consumption growth and dividend growth, we keep ρ_{ci} relatively small (equal to 0.1) to insure positive semi-definiteness of Ω . Second, many parameter combinations are ruled out by the requirement that the price-dividend ratio be finite. Thus, risk-aversion cannot be too low if we are to consider cases in which dividend growth grows more quickly than certain threshold amounts. We only consider parameter combinations that deliver both a positive semi-definite variance-covariance matrix and a finite price-dividend ratio. The table reports results for which γ is set to 5, δ is set to 0.99, $\sigma_i/\sigma_c = 1, 2, 4$, $\mu_i/\mu_c = 0.85, 1.5$, $\rho_{ci} = 0.1$, $\rho_{id} = 0.9$.

Table 10 shows a range of cases in which the joint distribution deviates considerably from normality and yet the pricing errors associated with erroneously using M_t^c to price assets in place of M_t^i are, to numerical accuracy, zero. In one case, the marginal distribution for dividend growth is close to normal, but those for stockholder and aggregate consumption growth deviate considerably from normality. For example, the kurtosis for Δc_t is often greater than 11, and the skewness greater than 4. For such cases, the values of γ_c and δ_c that drive the pricing errors to zero range but are typically not close to the true preference parameters for asset-holder i . The parameter γ_c is much larger than the true γ when assetholder consumption growth is much more volatile than aggregate consumption growth or when it is not highly correlated with it, as suggested by (17). Also, when $\text{Cov}(\Delta c, \Delta d) = \sigma_{cd}$ is parameterize to be negative, γ_c takes on negative values. This is similar to the normal case

(13), where the expression in (13) collapses to $\gamma_c = \gamma\sigma_{id}/\sigma_{cd}$, so that γ_c is negative when σ_{cd} is negative. Again, however, in every case reported in Table 4, and regardless of the values taken by γ_c and δ_c , the pricing errors for the stock return and risk-free rate are always numerically zero. This was also true of several thousand other parameter combinations.

Figure 6 provides a graphical description of two of the perturbed densities, plotted along with the leading normal, that created the output in Table 10. Notice that the shapes can differ considerably from Gaussian and yet we can still find values for γ_c and δ_c for which M_t^c prices assets just as well as the true kernel based on assetholder consumption. The densities in the left-hand column of Figure 2 are bimodal for Δc^i and Δd , while aggregate consumption is close to normal. This is similar to the example above (Figure 5), which does deliver large pricing errors using M_t^c , but unlike that case the negative mass points are much smaller relative to the positive mass points. In Figure 5, assetholder consumption growth has a higher mean than aggregate consumption growth, whereas in Figure 5 it has about the same mean. The densities in the right-hand column of Figure 6 are close to normal for Δc^i and Δd , while the density of aggregate consumption has skewness of about 4 and kurtosis around 11, strongly non-normal. These examples illustrate an important point, namely that there are many ways to depart from normality in limited participation/incomplete markets models that do not deliver an explanation of the empirical facts documented in Section 2.

3.3.1 Larger Cross-Sections of Returns

The analysis above was conducted for the case of two asset returns, a single risky asset and a risk-free asset. To evaluate the pricing errors for a larger cross-section of returns, we consider simple models of N assets, indexed by j , whose dividend processes take the form

$$\Delta d^j = \lambda^j \Delta c_t^i + \varepsilon_t^j, \quad j = 1, \dots, N,$$

where ε_t^j is an i.i.d. shock uncorrelated with Δc_t^i . In analogy to the two-asset case above, the vector of variables $y_t = (\Delta c_t, \Delta c_t^i, \Delta d_t^1, \dots, \Delta d_t^N)'$ is assumed to be i.i.d. Since the log stochastic discount factor, m^i , is linear in the assetholder's consumption growth, the "leverage" parameter λ^j controls the covariance of each asset return with the log stochastic discount factor, and ε_t^j controls the variance of individual risky returns. Assets on the log mean-variance efficient frontier (i.e., those that are perfectly correlated with m^i) have shocks ε_t^j with zero variance. The farther an asset return is from the log mean-variance efficient frontier, the larger the variance of ε_t^j . By varying λ^j across assets, we create a spread in the covariance of returns with stockholder consumption growth, and therefore a spread in risk-premia. The purpose of this exercise is not to create a realistic cross-section of asset returns, but rather to investigate the results in an overidentified setting with more moment conditions than parameters to be estimated.

We use the same procedure as for the two-asset case to describe the joint density, $h(y)$, of aggregate consumption, asset-holder consumption and the N asset returns. We calibrate the leading normal for $N = 8$ artificial assets, including a risk-free return, with $\lambda^j = \varepsilon_t^j = 0$, and a mean-variance efficient return that is perfectly correlated with the log stochastic discount factor, $\lambda^j = 1$ and $\varepsilon_t^j = 0$. The six other asset returns are generated by a grid of values of λ^j and $\text{Var}(\varepsilon_t^j)$, and the equilibrium returns R_{t+1}^j are computed as described in the previous subsection for to two-asset return case.

We search numerically for values for δ_c and γ_c to minimize the quadratic form $\mathbf{g}(\gamma_c, \delta_c) \equiv \mathbf{w}'(\gamma_c, \delta_c) \mathbf{W} \mathbf{w}(\gamma_c, \delta_c)$, where $\mathbf{w}(\gamma_c, \delta_c)$ is the (8×1) vector of pricing errors for each asset (i.e., $w_j(\gamma_c, \delta_c) = E \left(\delta_c \left(\frac{C_t}{C_{t-1}} \right)^{-\gamma_c} R_t^j - 1 \right)$ for $j = 1, \dots, 8$) and \mathbf{W} is the 8×8 identity matrix. As before, we consider a large number of possible distributions $h(y)$ by considering different parameter combinations in $a(y)^2$. Calibration of other parameters remains the same as above for the two-asset case.

Table 11 presents the square root of the average-squared pricing error (RMSE), as a fraction of the cross-sectional mean of the average returns, for several parameter assumptions. Instead of presenting the results for thousands of distributional assumptions (corresponding to hundreds of possible combinations of the nine parameters in $a(y)^2$), we present the maximum and average RMSE (as a fraction of the cross-sectional average mean return) obtained over a large grid search of distributional parameters. We also present these figures for the special case in which all variables are jointly lognormally distributed. The results are again that the average pricing errors that arise from using aggregate consumption growth in place of asset-holder consumption growth are often very small, indeed close to zero. For example, when all variables are jointly lognormally distributed, the RMSE is 0.02% as a fraction of the cross-sectional average mean return, whereas in the data they are 35% (Table 1). These results echo the findings for the two-asset case: when variables are jointly lognormally distributed, a representative-agent pricing kernel based on per capita aggregate consumption can often be found that has the same asset-pricing implications as the true pricing kernel, although estimates of γ and δ will not correspond to any individual's true risk-aversion or time discount factor.

When we consider perturbations from the lognormal case, the typical result is again a small RMSE. For example, when the covariance between aggregate consumption growth and dividend growth is set to a small negative number (to match the empirical covariance in annual data from Lettau and Ludvigson (2004)), the maximum RMSE, as a fraction of the cross-sectional mean return, is 0.58%, and the average is 0.03%. Again this should be compared with the 35% in the data. By contrast, when the covariance is set to a small positive number (the negative of the point estimate in the data), we find a small number of cases in

which the pricing error as a fraction of the cross-sectional average return is as large as 10%. But these cases are relatively rare and occur in less than 0.2% of the parameter combinations. Most models that use the wrong pricing kernel based on aggregate consumption deliver tiny pricing errors even when aggregate consumption has a low correlation with asset-holder consumption and only half its volatility. These results echo the findings reported above for the two-asset case.

4 Conclusion

The asset pricing literature has delivered a number of prominent new theories in recent years designed specifically to remedy shortcomings of the standard consumption-based asset pricing model. In this paper we emphasize one shortcoming of the standard model that has received little attention but that nevertheless provides a margin upon which the model fails overwhelmingly: its inability to satisfy the unconditional moment restrictions implied by theory. This failure is quantitatively large and present even when the range of parameters for risk aversion and time preference is left unrestricted and chosen to maximize the model's chance of success. We argue that these empirical facts constitutes a puzzle for the standard model that is at least as damning other, more well known, puzzles commonly emphasized when studying calibrated models.

Are prominent modifications to the standard model capable of explaining these phenomena? If so, then in these models use of the standard pricing kernel to explain asset returns should generate large unconditional asset pricing errors, as in the data. Alas, we find that new pricing kernels do not necessarily generate new pricing implications. Instead, we show that if asset return and consumption data were in fact generated by leading asset pricing models, then parameters for risk aversion and time preference could almost certainly be found that imply the standard model has the same explanatory power in tests of unconditional asset pricing restrictions as those models currently at the forefront of theoretical asset pricing. This is true both for explaining the behavior of a single risky asset and a risk-free asset, and for explaining larger cross-sections of risky returns. Moreover, some leading models imply that the standard model is equally capable of explaining asset returns even when it is based both on the wrong consumption measure and on the wrong underlying preferences. We show that leading asset pricing models *can*, in many cases, explain why an econometrician obtains implausibly high estimates of δ and γ when fitting data to the standard consumption-based model. But they cannot explain why the standard model fails so resoundingly to satisfy the most basic unconditional moment restrictions implied by theory. A complete explanation of aggregate stock market behavior must account for these empirical regularities.

Failure to account for these empirical regularities cannot be uncovered by studying calibrated models or by estimation procedures that rely solely on a model's first-order conditions for identification. That is because the first-order conditions of any model are not a complete description of the joint distribution of asset returns and aggregate quantities. But an econometrician who observes this joint distribution in the data can use it to ask whether its key properties are matched by the simulated data of theoretical models.

Intuitively, how is it that asset pricing models capable of explaining the equity premium puzzle, and the host of other consumption-based asset pricing puzzles in the literature, are incapable of explaining the large unconditional Euler equation errors of the standard consumption-based model? In thinking about this, it is helpful to consider a simple example. We know that the equity premium puzzle can be “solved” by taking the standard consumption-based model and applying sufficiently high risk-aversion (Mehra and Prescott (1985)).¹⁶ The difficulty with this resolution of the puzzle is that, in order to show that high risk-aversion delivers the right equity premium as an equilibrium outcome, the resulting equilibrium returns must be derived from theoretical Euler equations that are exactly satisfied. To the extent that these Euler equations are not satisfied in historical data, such a resolution would seem to rest on a fundamental misspecification of the joint behavior of asset returns and aggregate quantities.

What types of changes might bring asset pricing models more in line with the data along this dimension? We provide one example of a limited participation/incomplete markets model that can explain the large unconditional Euler equation errors generated by the standard model. In this example, consumption of asetholders is considerably more volatile than aggregate consumption and weakly correlated with it, whereas it is highly correlated with the risky asset return. But it is not enough for asetholder consumption to merely behave differently from aggregate consumption. In the example we provide, the distribution of aggregate consumption is close to normal, but the densities of asetholder consumption growth and returns appear highly non-normal and strongly bimodal. This suggests that careful attention to the joint properties of the pricing kernel, aggregate consumption, and returns is crucial for explaining these empirical regularities in frictionless models. Classes of economic models with endogenously distorted beliefs, as surveyed in the work of Hansen and Sargent (2000) or illustrated in the learning model of Cogley and Sargent (2004), also suggest interesting possibilities for explaining these phenomena. In such models, beliefs are distorted away from what a model of rational expectations would impose, so asset return volatility can be driven by fluctuations in beliefs not necessarily highly correlated with consumption. Other

¹⁶This creates a risk-free rate puzzle and the resulting model has a number of other shortcomings, hence leading asset pricing models depart from the standard model in more substantive ways. But for the sake of building intuition, we may assume that equity premium puzzle is all we care about.

candidates include any modifications to the standard model that would make unconditional Euler equations more difficult to satisfy with equality, especially in recessions. Possibilities include binding restrictions on the ability to trade and smooth consumption, such as borrowing constraints, short-sales constraints, and transactions costs (e.g., Luttmer (1996); He and Modest (1995); Heaton and Lucas (1996, 1997); Ludvigson (1999)). An important area for future research will be to determine whether such modifications are capable of delivering the empirical facts, once introduced into plausibly calibrated economic models with empirically credible frictions.

5 Appendix

1. Data Description

This appendix describes the data. The sources and description of each data series we use are listed below.

CONSUMPTION

Consumption is measured as expenditures on nondurables and services, excluding shoes and clothing. The quarterly data are seasonally adjusted at annual rates, in billions of chain-weighted 1996 dollars. The components are chain-weighted together, and this series is scaled up so that the sample mean matches the sample mean of total personal consumption expenditures. Our source is the U.S. Department of Commerce, Bureau of Economic Analysis.

POPULATION

A measure of population is created by dividing real total disposable income by real per capita disposable income. Consumption, is in per capita terms. Our source is the Bureau of Economic Analysis.

PRICE DEFLATOR

Real asset returns are deflated by the implicit chain-type price deflator (1996=100) given for the consumption measure described above. Our source is the U.S. Department of Commerce, Bureau of Economic Analysis.

ASSET RETURNS

- Three-Month Treasury Bill Rate: secondary market, averages of business days, discount basis%; Source: H.15 Release – Federal Reserve Board of Governors.
- Six size/book-market returns: Six portfolios, monthly returns from July 1926-December 2003. The portfolios, which are constructed at the end of each June, are the intersections of 2 portfolios formed on size (market equity, ME) and 3 portfolios formed on the ratio of book equity to market equity (BE/ME). The size breakpoint for year t is the median NYSE market equity at the end of June of year t . BE/ME for June of year t is the book equity for the last fiscal year end in $t-1$ divided by ME for December of $t-1$. The BE/ME breakpoints are the 30th and 70th NYSE percentiles. Source: Kenneth French's homepage, http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.
- The stock market return is the Center for Research and Security Prices (CRSP) value-weighted stock market return. Our source is the Center for Research in Security Prices.

2. Detailed Description of Models

The utility function in the CC and MSV models take the form

$$U = E \left\{ \sum_{t=0}^{\infty} \delta^t \frac{(C_t^i - X_t^i)^{1-\gamma} - 1}{1-\gamma} \right\}, \quad \gamma > 0 \quad (24)$$

where C_t^i is individual consumption and X_t is habit level which they assume to be a function of aggregate consumption, and δ is the subjective discount factor. In equilibrium, identical agents choose the same level of consumption, so C_t^i is equal to aggregate consumption, C_t . CC define the surplus consumption ratio

$$S_t \equiv \frac{C_t - X_t}{C_t} < 1,$$

and model its log process as evolving according to a heteroskedastic first-order autoregressive process (where as before lowercase letters denote log variables):

$$s_{t+1} = (1 - \phi) \bar{s} + \phi s_t + \lambda(s_t) (c_{t+1} - c_t - g),$$

where ϕ , g , and \bar{s} are parameters. $\lambda(s_t)$ is the so-called sensitivity function that CC choose to satisfy three conditions: (1) the risk-free rate is constant, (2) habit is predetermined at steady state, and (3) habit moves nonnegatively with consumption everywhere. We refer the reader to the CC paper for the specific functional form of $\lambda(s_t)$. The stochastic discount factor in the CC model is given by

$$M_{t+1}^{CC} = \delta \left(\frac{C_{t+1}}{C_t} \frac{S_{t+1}}{S_t} \right)^{-\gamma}.$$

In all of the models considered here, the return on a risk-free asset whose value is known with certainty at time t is given by

$$R_{t+1}^f \equiv (E_t [M_{t+1}])^{-1},$$

where M_{t+1} is the pricing kernel of whichever model we are considering.

MSV model the behavior of Y_t , the inverse surplus consumption ratio:

$$Y_t = \frac{1}{1 - (X_t/C_t)} > 1.$$

Following Campbell and Cochrane (1999), MSV assume that Y_t follows a mean-reverting process, perfectly negatively correlated with innovations in consumption growth:

$$\Delta Y_t = k (\bar{Y} - Y_t) - \alpha (Y_t - \lambda) (\Delta c_t - E_{t-1} \Delta c_t),$$

where \bar{Y} is the long-run mean of Y and k , α , and λ are parameters, calibrated as in MSV. Here $\Delta c_t \equiv \log(C_t/C_{t-1})$, which they assume it follows an i.i.d. process

$$\Delta c_t = \mu + \sigma v_t,$$

where v_t is a normally distributed i.i.d. shock. The stochastic discount factor in the MSV model is

$$M_{t+1}^{MSV} = \delta \left(\frac{C_{t+1}}{C_t} \frac{Y_t}{Y_{t+1}} \right)^{-\gamma}.$$

Since the MSV model is a representative-agent model, we modify it in order to study the role of limited participation. Assume that asset prices are determined by the framework above, where a valid stochastic discount factor is a function of any stockholder's consumption C_t^i and stockholder's habit X_t^i . The process for stockholder consumption is the same as in MSV, described above, but now with i subscripts:

$$\Delta c_t^i = \mu_i + \sigma_i v_t^i,$$

where v_t^i is a normally distributed i.i.d. shock. Aggregate consumption is assumed to follow a separate process given by

$$\Delta c_t = \mu_c + \sigma_c v_t^c,$$

with v_t^c a normally distributed i.i.d. shock. We analyze the results over a range of cases for the correlation between v_t^i and v_t^c , and their relative volatilities σ_i/σ_c .

For the representative stockholder, we model the first difference of Y_t^i as in MSV:

$$\Delta Y_t^i = k \left(\bar{Y}^i - Y_t^i \right) - \alpha \left(Y_t^i - \lambda \right) \left(\Delta c_t^i - E_{t-1} \Delta c_t^i \right),$$

and compute equilibrium asset returns based on the stochastic discount factor $M_{t+1}^{MSV^i} = \delta \left(C_{t+1}^i / C_t^i \right)^{-\gamma} \left(Y_t^i / Y_{t+1}^i \right)^{-\gamma}$. As before, this is straightforward to do using the analytical solutions provided in MSV.

Next, we compute two types of unconditional pricing errors. First, we compute the pricing errors generated from erroneously using aggregate consumption in the pricing kernel in place of assetholder consumption. That is, we compute the pricing errors that arise from using $M_{t+1}^{ch} \equiv \delta^c \left(C_{t+1}^c / C_t^c \right)^{-\gamma^c} \left(Y_t^c / Y_{t+1}^c \right)^{-\gamma^c}$ in place of $M_{t+1}^{MSV^i}$ to price assets, where δ^c and γ^c are chosen freely to fit the data, and where Y_t^c follows the process

$$\Delta Y_t^c = k \left(\bar{Y}^c - Y_t^c \right) - \alpha \left(Y_t^c - \lambda \right) \left(\Delta c_t - E_{t-1} \Delta c_t \right).$$

With the exception of α , all parameters are set as in MSV. The parameter α is set to keep the mean return on the aggregate wealth portfolio the same as in MSV. Thus, if $\sigma_i/\sigma_c = 2$, the value of α in MSV is divided by two.

To model multiple risky securities, MSV model the share of aggregate consumption that each asset produces,

$$s_t^j = \frac{D_t^j}{C_t} \quad \text{for } j = 1, \dots, n,$$

where n represents the total number of risky financial assets paying a dividend D . MSV assume that these shares are bounded, mean-reverting and evolve according to

$$\Delta s_t^j = \phi^j (\bar{s}^j - s_t^j) + s_t^j \boldsymbol{\sigma}(s_t) \boldsymbol{\epsilon}_t,$$

where $\boldsymbol{\sigma}(s_j)$ is an N -dimensional row vector of volatilities and $\boldsymbol{\epsilon}_t$ is an N -dimensional column vector of standard normal random variables, and ϕ^j and \bar{s}^j are parameters. ($N \leq n + 1$ because MSV allow for other sources of income, e.g., labor income, that support consumption.) Cross-sectional variation in unconditional mean returns across risky securities in this model is governed by cross-sectional variation in the covariance between shares and aggregate consumption growth: $\text{Cov}\left(\frac{\Delta s_t^j}{s_t^j}, \frac{\Delta c_t}{c_t}\right)$, for $j = 1, \dots, n$. This in turn is determined by cross-sectional variation in ϕ^j , \bar{s}^j and $\boldsymbol{\sigma}(s_j)$. We create n artificial risky securities using an evenly spaced grid of values for these parameters. The values of ϕ^j lie on a grid between 0 and 1, and the values of $\bar{s}^j \in [0, 1)$ lie on a grid such that the sum over all j is unity. The parametric process for $\boldsymbol{\sigma}(s_j)$ follows the specification in MSV in which the volatilities depend on a N -dimensional vector of parameters \mathbf{v}^j as well as the individual share processes

$$\boldsymbol{\sigma}(s_j) = \mathbf{v}^j - \sum_{k=0}^n s_t^k \mathbf{v}^k.$$

We choose the parameters ϕ^j , \bar{s}^j , and \mathbf{v}^j , to generate a spread in average returns across assets. In analogy to the empirical exercise (Panel B of Table 1), we do this for $n = 6$ risky securities plus the aggregate wealth portfolio return and the risk-free for a total of 8 asset returns.

Closed-form solutions are not available for the individual risky securities, but MSV show that equilibrium price-dividend ratios on the risky assets are given by the approximate relation

$$\frac{P_t^j}{D_t^j} \approx a_0^j + a_1^j S_t + a_2^j \frac{\bar{s}^j}{s_t^j} + a_3^j \frac{\bar{s}^j}{s_t^i} S_t, \quad (25)$$

where $S_t \equiv 1/Y_t^i$ and where Y_t^i again denotes the inverse surplus ratio of an individual assetholder indexed by i , which should not be confused with the indexation by j , which denotes a security. The parameters a_0^j , a_1^j , a_2^j , and a_3^j are all defined in terms of the other parameters above. Using these solutions for individual price-dividend ratios, we create a cross-section of equilibrium risky securities using

$$R_{t+1}^i = \left(\frac{P_{t+1}^j/D_{t+1}^j + 1}{P_t^j/D_t^j} \right) \exp(\Delta d_{t+1}^j). \quad (26)$$

Bansal and Yaron (2004) consider a representative-agent who maximizes utility given by recursive preferences of Epstein and Zin (1989, 1991) and Weil (1989). The utility function to be maximized takes the form

$$U = E \left\{ \sum_{t=0}^{\infty} \delta^t \left\{ (1 - \delta) C_t^{\frac{1-\gamma}{\alpha}} + \delta (E_t U_{t+1}^{1-\gamma})^{\frac{1}{\alpha}} \right\}^{\frac{\alpha}{1-\gamma}} \right\}, \quad (27)$$

where $\alpha \equiv (1 - \gamma) / (1 - 1/\psi)$, ψ is the intertemporal elasticity of substitution in consumption (IES), γ is the coefficient of relative risk aversion, and δ is the subjective discount factor. The stochastic discount factor under Epstein-Zin-Weil utility used in BY takes the form

$$M_{t+1}^{BY} = \left(\delta \left(\frac{C_{t+1}}{C_t} \right)^{-\frac{1}{\psi}} \right)^{\alpha} R_{w,t+1}^{\alpha-1}, \quad (28)$$

where $R_{w,t+1}$ is the simple gross return on the aggregate wealth portfolio, which pays a dividend equal to aggregate consumption, C_t .

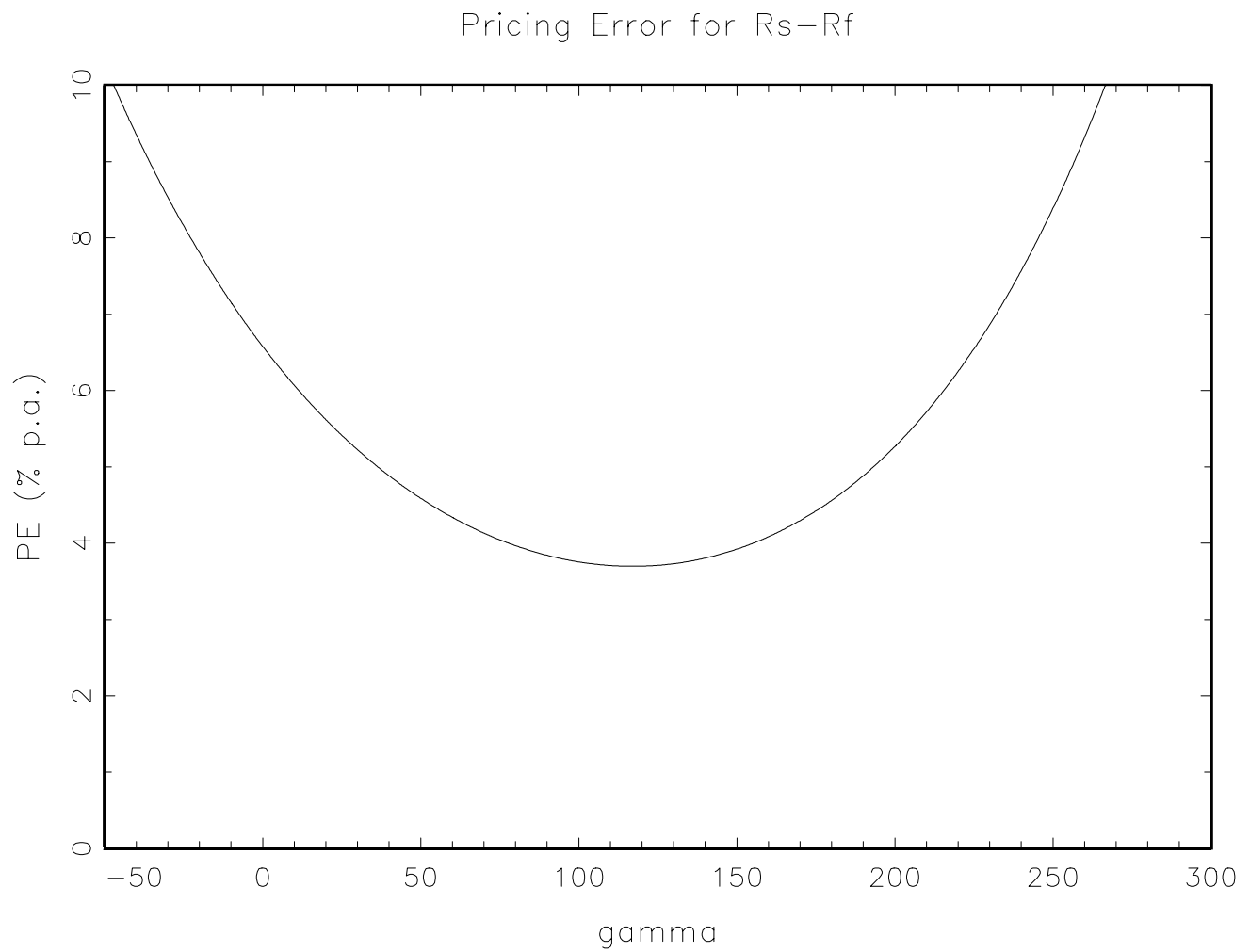
References

- BANSAL, R., AND A. YARON (2004): “Risks for the Long-Run: A Potential Resolution of Asset Pricing Puzzles,” *Journal of Finance*, 59(4), 1481–1509.
- CAMPBELL, J. Y. (2003): “Consumption-Based Asset Pricing,” in *Handbook of the Economics of Finance* forthcoming, ed. by G. Constantinides, M. Harris, and R. Stulz. North-Holland, Amsterdam.
- CAMPBELL, J. Y., AND J. H. COCHRANE (1999): “By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior,” *Journal of Political Economy*, 107, 205–251.
- COGLEY, T., AND T. J. SARGENT (2004): “The Market Price of Risk and the Equity Premium: A Legacy of the Great Depression?,” Unpublished paper, New York University.
- CONSTANTINIDES, G. M., AND D. DUFFIE (1996): “Asset Pricing With Heterogeneous Consumers,” *Journal of Political Economy*, 104, 219–40.
- EPSTEIN, L., AND S. ZIN (1989): “Substitution Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework,” *Econometrica*, 57, 937–968.
- (1991): “Substitution, Risk Aversion, and the Temporal behavior of Consumption and Asset Returns: An Empirical Investigation,” *Journal of Political Economy*, 99, 555–576.
- FAMA, E. F., AND K. R. FRENCH (1992): “The Cross-Section of Expected Returns,” *Journal of Finance*, 47, 427–465.
- FERSON, W. E., AND G. M. CONSTANTINIDES (1991): “Habit Persistence and Durability in Aggregate Consumption,” *Journal of Financial Economics*, 29, 199–240.
- GALLANT, A. R., AND G. TAUCHEN (1989): “Seminonparametric Estimation of Conditionally Constrained Heterogeneous Processes: Asset Pricing Applications,” *Econometrica*, 57(5), 1091–1120.
- GORDON, S., AND P. ST-AMOUR (2004): “Asset Returns and State Dependent Risk Preferences,” *Journal of Business and Economic Statistics*, 22, 241–252.
- GUVENEN, M. F. (2003): “A Parsimonious Macroeconomic Model for Asset Pricing: Habit Formation or Cross-sectional Heterogeneity?,” Rochester Center for Economic Research Working Paper No. 499.
- HANSEN, L. P. (1982): “Large Sample Properties of Generalized Methods of Moments Estimators,” *Econometrica*, 50, 1029–54.
- HANSEN, L. P., AND R. JAGANNATHAN (1991): “Restrictions on Intertemporal Marginal Rates of Substitution Implied by Asset Returns,” *Journal of Political Economy*, 99, 225–262.

- HANSEN, L. P., AND T. J. SARGENT (2000): "Wanting Robustness in Macroeconomics," Unpublished Paper, New York University.
- HANSEN, L. P., AND K. SINGLETON (1982): "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica*, 50(5), 1269–86.
- HE, H., AND D. M. MODEST (1995): "Portfolio Choice and Asset Prices: The Importance of Entrepreneurial Risk," *Journal of Finance*, 103, 94–117.
- HEATON, J., AND D. LUCAS (1996): "Evaluating the Effects of Incomplete Markets on Risk Sharing and Asset Pricing," *Journal of Political Economy*, 104(3), 443–87.
- (1997): "Market Frictions, Savings Behavior, and Portfolio Choice," *Macroeconomic Dynamics*, 1(1), 76–101.
- JAGANNATHAN, R., AND Y. WANG (2004): "Consumption Risk and the Cost of Equity Capital," Unpublished paper, Kellogg School of Management, Northwestern University.
- KOCHERLAKOTA, N. R. (1990): "Disentangling the Coefficient of Relative Risk Aversion from the Elasticity of Intertemporal Substitution: An Irrelevance Result," *Journal of Finance*, 45(1), 175–190.
- KREBS, T. (2004): "Testable Implications of Consumption-Based Asset Pricing Models with Incomplete Markets," *Journal of Mathematical Economics*, 40, 191–206.
- LETTAU, M., AND S. C. LUDVIGSON (2004): "Expected Returns and Expected Dividend Growth," *Journal of Financial Economics*, forthcoming.
- LUDVIGSON, S. C. (1999): "Consumption and Credit: A Model of Time-Varying Liquidity Constraints," *Review of Economics and Statistics*, 81, 434–447.
- LUTTMER, E. G. J. (1996): "Asset Pricing in Economies with Frictions," *Econometrica*, 64(6), 1439–1467.
- MARDIA, K. (1970): "Measures of Multivariate Skewness and Kurtosis with Applications," *Biometrika*, 57, 519–530.
- MEHRA, R., AND E. C. PRESCOTT (1985): "The Equity Premium Puzzle," *Journal of Monetary Economics*, 15, 145–161.
- MELINO, A., AND A. YANG (2003): "State-dependent preferences can explain the equity premium puzzle," *Review of Economic Dynamics*, 6(4), 806–830.
- MENZLY, L., T. SANTOS, AND P. VERONESI (2004): "Understanding Predictability," *Journal of Political Economy*, 112(1), 1–47.

WEIL, P. (1989): "The Equity Premium Puzzle and the Risk-Free Rate Puzzle," *Journal of Monetary Economics*, 24(3), 401–421.

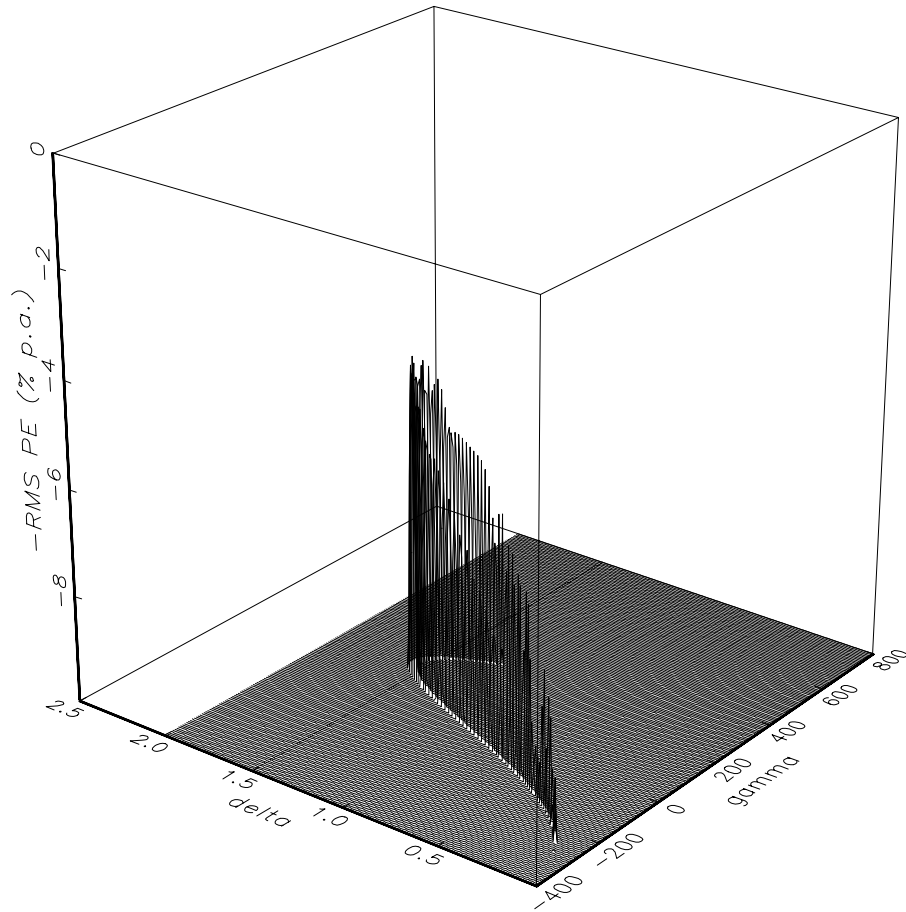
Figure 1: Pricing Errors for CRRA Preferences



Notes: The figure plots the annualized pricing error $E[(C_t/C_{t-1})^{-\gamma}(R_t^s - R_t^f)]$ (in %) as a function of γ .

Figure 2: Pricing Errors for CRRA Preferences

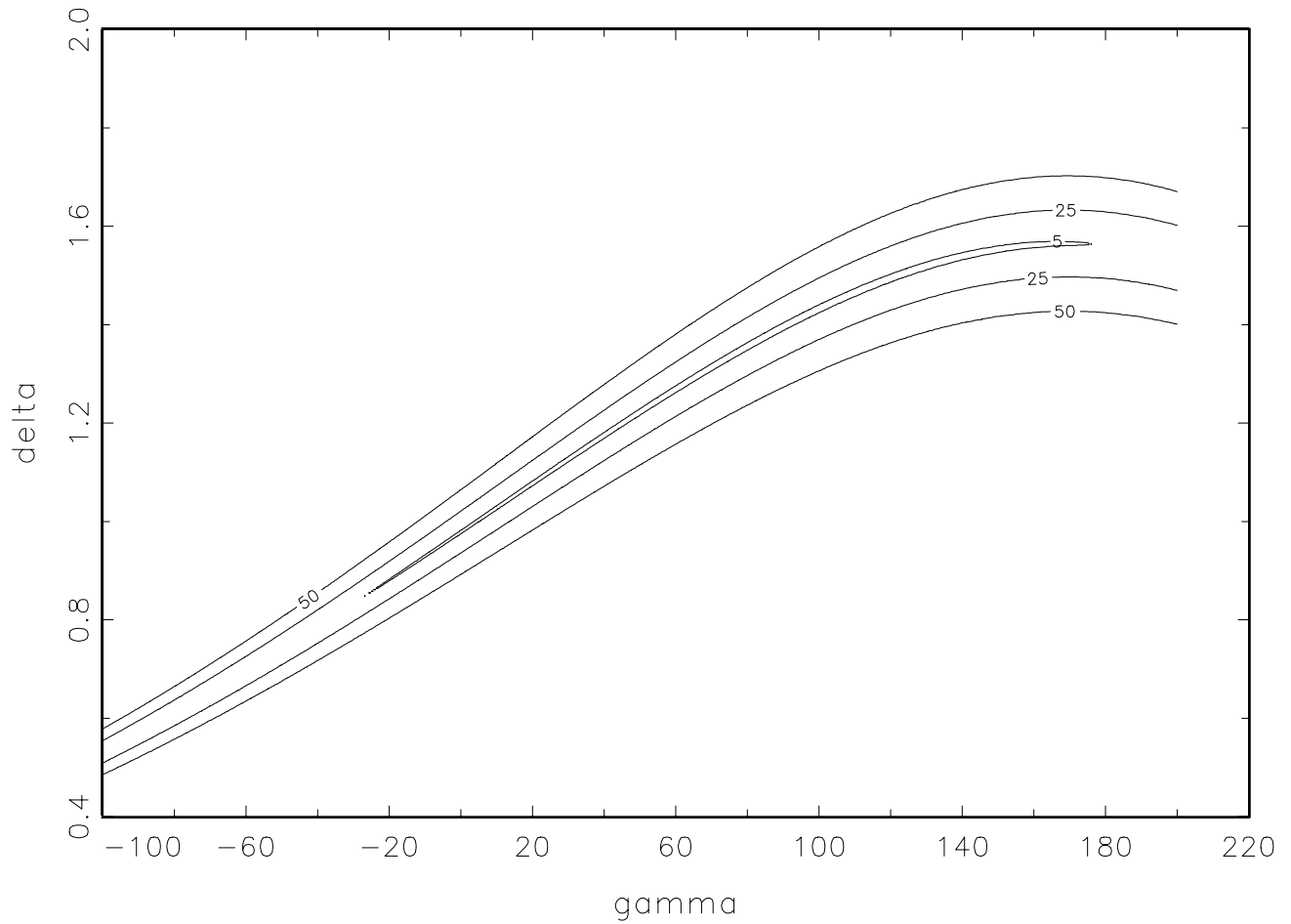
Root-Mean-Squared Pricing Errors of R_s and R_f



Notes: The figures plot the negative of the annualized root-mean-squared pricing errors $[E[\delta(C_t/C_{t-1})^{-\gamma} R_t^s - 1]^2 + E[\delta(C_t/C_{t-1})^{-\gamma} R_t^f - 1]^2]^{1/2}$ (in %) implied by CRRA preference as a function of γ and δ . Pricing errors that exceed 10% are assigned a value of 10%.

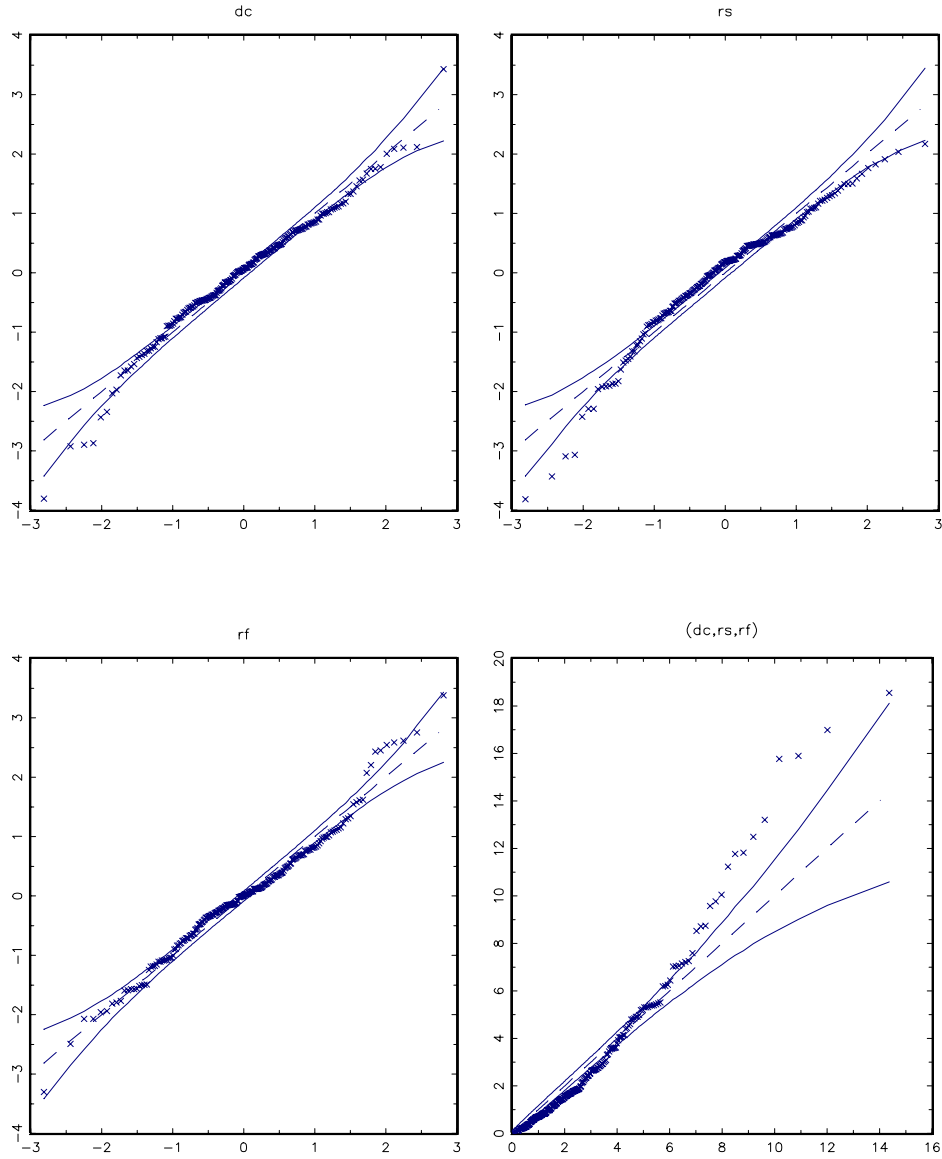
Figure 3: Pricing Errors: Contour Plots

Contour Plot of Pricing Errors: RMSE (% p.a.)



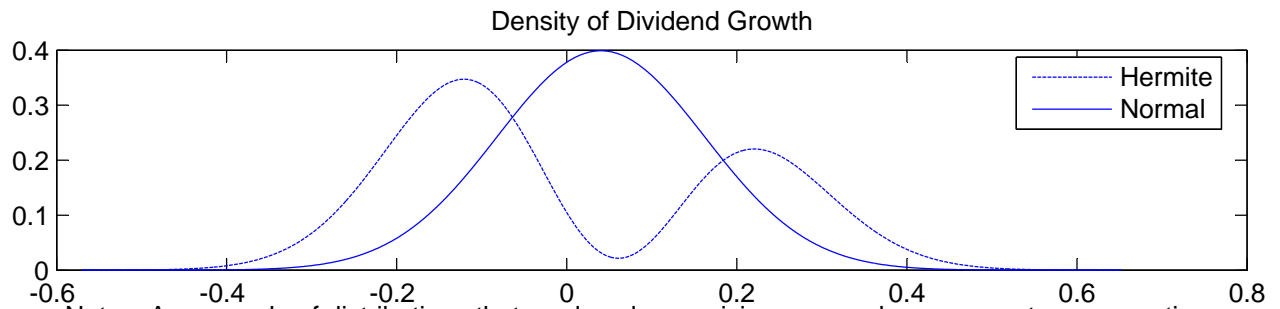
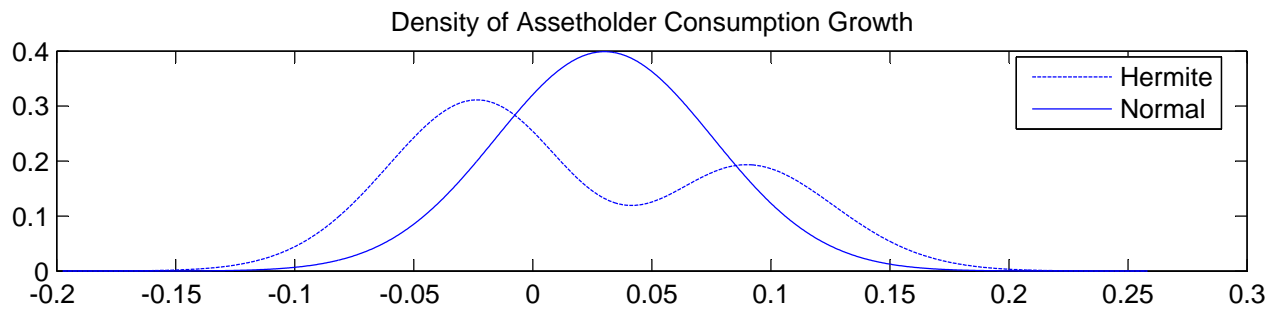
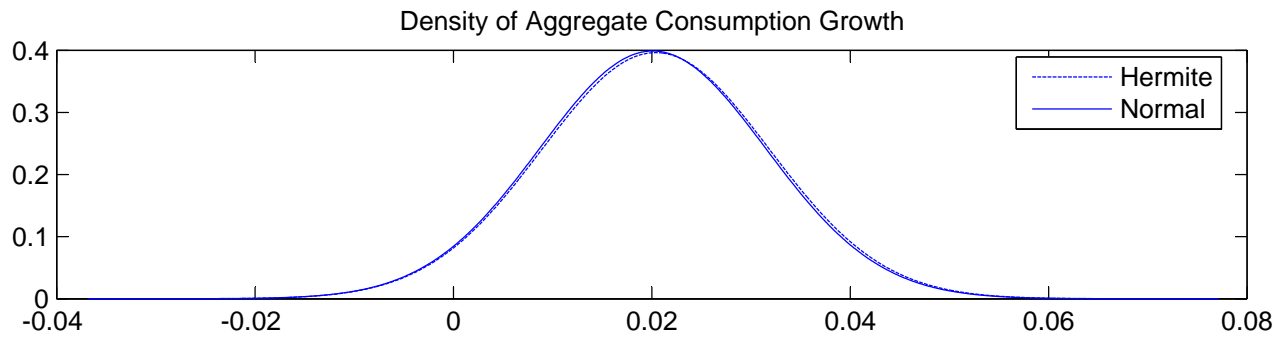
Notes: This figure shows contour plots of the pricing errors depicted in Figure 1.

Figure 4: QQ Plots



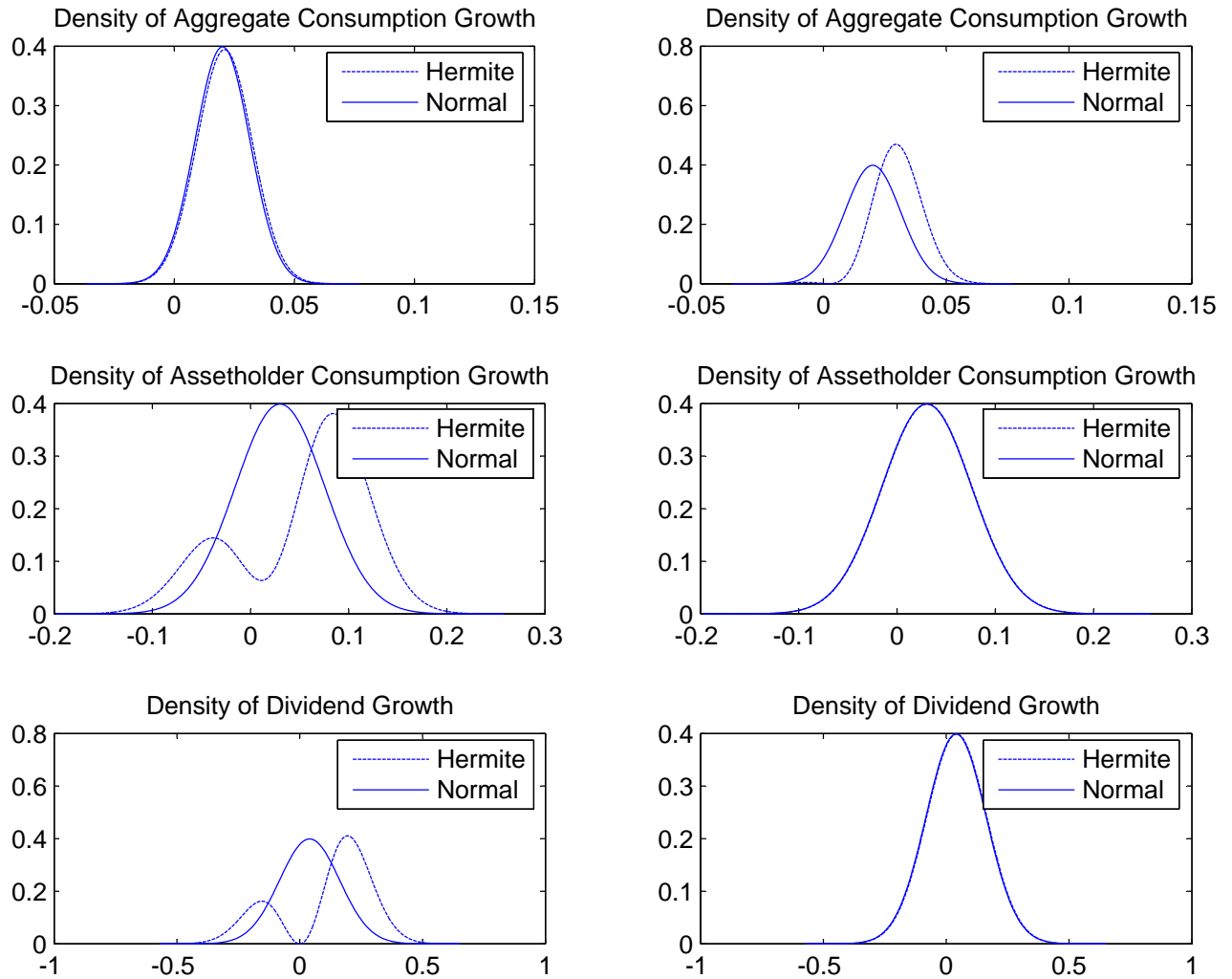
Notes: This figure shows quantile-quantile (QQ) plots of the logs of consumption growth, stock returns, and the riskfree rate. Each panel plots the sample quantiles (on the y -axis) against the quantiles of a given distribution (on the x -axis) as well pointwise 5% and 95% bands. The univariate QQ plots (denoted Δc , r_s and r_f) compare quantiles of the sample distributions to those of a normal distribution. The multivariate QQ plot for the joint distribution of Δc , r_s and r_f shows the quantiles of the squared Mahalanobis distances against those of a χ_3^2 distribution. The squared Mahalanobis distance M_t for a p -dimensional multivariate distribution \mathbf{x}_t with mean $\boldsymbol{\mu}_x$ and variance-covariance matrix \mathbf{V} is defined as $M_t = (\mathbf{x}_t - \boldsymbol{\mu}_x)' \mathbf{V}^{-1} (\mathbf{x}_t - \boldsymbol{\mu}_x)$. Under the null hypothesis that Δc , r_s and r_f are jointly normally distributed, M_t has a χ_p^2 distribution.

Figure 5



Notes: An example of distributions that produce large pricing errors when aggregate consumption and returns are fitted to a power utility model.

Figure 6



Notes: Plots of marginal densities for two Hermite parameter configurations.

Table 1: Pricing Error with CRRA Preferences

Panel A: CRSP-VW Return and Riskfree Rate						
ω_s	ω_f	δ_c	γ_c	PE(R^f) (in %)	PE(R^s) (in %)	RMSE/ $E[\mathbf{R}_t]$
1	1	1.411	89.780	-2.73	2.69	0.60
1000	1	1.405	90.688	-5.38	0.00	0.84
1	1000	1.425	90.881	0.00	5.38	0.84
Excluding Periods with low Consumption Growth						
1	1	2.555	326.117	-0.73	0.73	0.16
1000	1	2.550	366.128	-1.46	0.00	0.23
1	1000	1.478	83.103	0.00	1.46	0.23
Panel B: CRSP-VW Return, Riskfree Rate, and 6 FF Portfolios						
δ	γ	RMSE	RMSE/ $E[\mathbf{R}_t]$			
1.39	87.18	3.05	0.35			
Excluding Periods with low Consumption Growth						
2.58	356.07	1.94	0.23			

Notes: This table reports the minimized annualized postwar data pricing error for CRRA preferences. The preference parameters δ_c and γ_c are chosen to minimize the mean square pricing error for different sets of returns: $\min_{\delta, \gamma} [g(\delta, \gamma)' W g(\delta, \gamma)]$ where $g(\delta, \gamma) = E[\delta(C_t/C_{t-1})^{-\gamma} \mathbf{R}_t - 1]$ and W is a weighting matrix. In Panel A, \mathbf{R} includes R^s and R^f and the weighting matrix is $W = [\omega_s \ 0, 0 \ \omega_f]$. In Panel B, \mathbf{R} includes the six baseline Fama-French portfolios and $W = I$. The table also reports results when the periods with the lowest six consumption growth rates are eliminated. R^s is the CRSP-VW stock returns, R^f is the 3-month T-bill rate and C_t is real per-capita consumption of nondurables and services excluding shoes and clothing. PE(R^s) and PE(R^f) denote the average pricing errors $E[\delta_c(C_t/C_{t-1})^{-\gamma_c} R_t^s - 1]$ and $E[\delta_c(C_t/C_{t-1})^{-\gamma_c} R_t^f - 1]$. $E[\mathbf{R}_t]$ is the average of the mean returns of the assets under considerations. RMSE is the square root of the average squared pricing error. The data span the period 1951Q4 to 2002Q4.

Table 2: Low Consumption Growth Periods

Quarter	$C_t/C_{t-1} - 1$	R_t^s	R_t^f
1980Q02	-1.28	16.08	3.59
1990Q04	-0.87	8.75	2.16
1974Q01	-0.85	-1.26	2.37
1958Q01	-0.84	7.03	0.65
1960Q03	-0.64	-4.93	0.67
1953Q04	-0.60	7.87	0.47

Notes: This table reports consumption growth, the return of the CRSP-VW stock returns R^s and the 3-month T-bill rate R^f (all in in percent per quarter) in the six quarters of our sample with the lowest consumption growth rates. The consumption measure is real per-capita expenditures on nondurables and services excluding shoes and clothing. The data span the period 1951Q4 to 2002Q4.

Table 3: Tests of Joint Normality

	Skewness	p -value	Excess Kurtosis	p -value
Δc	0.43	0.01	1.44	0.00
r_s	0.89	0.00	1.46	0.00
r_f	0.22	0.20	1.06	0.00
$(\Delta c, r_s)$	1.11	0.00	2.72	0.00
$(\Delta c, r_f)$	1.05	0.00	3.12	0.00
(r_s, r_f)	1.06	0.00	2.84	0.00
$(\Delta c, r_s, r_f)$	1.54	0.00	4.64	0.00

Notes: This table reports multivariate skewness and kurtosis following Mardia (1970). Let \mathbf{x}_t be a p -dimensional random variable with mean $\boldsymbol{\mu}$ and variance-covariance matrix \mathbf{V} of sample size T . Multivariate skewness S and (excess) kurtosis K and asymptotic distributions are given by

$$S = \left(\frac{1}{T^2} \sum_{t=1}^T \sum_{s=1}^T g_{ts}^3 \right)^{1/2} \quad \frac{TS^2}{6} \sim \chi_{p(p+1)(p+2)/6}^2$$

$$K = \frac{1}{T} \sum_{t=1}^T g_{tt}^2 - p(p+2) \quad \frac{\sqrt{T}K}{\sqrt{(8p(p+2))}} \sim N(0, 1),$$

where $g_{ts} = (\mathbf{x}_t - \hat{\boldsymbol{\mu}})' \hat{\mathbf{V}}^{-1} (\mathbf{x}_s - \hat{\boldsymbol{\mu}})$ and $\hat{\boldsymbol{\mu}}$ and $\hat{\mathbf{V}}$ are sample estimates of $\boldsymbol{\mu}$ and \mathbf{V} . S and K are zero if \mathbf{x} is jointly normally distributed. If \mathbf{x} is univariate S and K are equivalent to the standard univariate definitions of skewness and kurtosis.

Table 4: Pricing Errors

Model	δ_c	γ_c	$E[\delta(C_t/C_{t-1})^{-\gamma} R_t^s - 1]$	$E[\delta(C_t/C_{t-1})^{-\gamma} R_t^f - 1]$
CC Habit	1.28	57.48	0.00%	0.00%
MSV Habit	1.71	30.64	0.00%	0.00%
BY LR Risk	0.93	48.97	0.00%	0.00%

Notes: This table reports the annualized pricing errors for stock returns R^s and the riskfree rate R^f from simulated data from Campbell and Cochrane's habit model (CC Habit), Menzly, Santos and Veronesi's habit model (MSV Habit) and Bansal and Yaron's long run risk model (BY LR Risk) for CRRA preferences. The preference parameters δ_c and γ_c are chosen to minimize the mean square pricing error $\min_{\delta, \gamma} [g(\delta, \gamma)' W g(\delta, \gamma)]$ where $g(\delta, \gamma) = E[\delta(C_t/C_{t-1})^{-\gamma} \mathbf{R}_t - 1]$. $\mathbf{R} = [R^s, R^f]'$ and $W = I$.

Table 5: Properties of Guvenen’s Model

Panel A: Consumption Growth					
	$C_t/C_{t-1} - 1$	$C_t^i/C_{t-1}^i - 1$	$C_t^n/C_{t-1}^n - 1$	R_t^s	R_t^f
Mean	0.01	0.02	0.00	1.31	0.64
Std. Dev.	2.04	4.53	0.83	7.30	1.69
Correlation	1.00	1.00	0.99	1.00	0.17
	1.00	1.00	0.98	0.99	0.17
	0.99	0.98	1.00	0.99	0.16
	1.00	0.99	0.99	1.00	0.19
	0.17	0.17	0.16	0.19	1.00
Panel B: Stochastic Discount Factor					
	$M_t^i(0.99, 2.00)$	$M_t^c(0.99, 2.00)$	$M_t^c(0.99, 4.49)$		
Mean	0.99	0.99	0.99		
Std. Dev.	0.09	0.04	0.09		
Correlation	1.00	1.00	1.00		
	1.00	1.00	1.00		
	1.00	1.00	1.00		
Panel C: Pricing Errors					
Consumption	(δ, γ)	$E[M_t(\delta, \gamma) R_t^s - 1]$	$E[M_t(\delta, \gamma) R_t^f - 1]$		
SH	(0.99, 2.00)	0.02%	0.02%		
AC	(0.99, 2.00)	0.39%	-0.34%		
AC	(0.99, 4.49)	0.00%	0.01%		

Notes: This table reports properties of Guevenen’s model. Panel A reports the properties of consumption growth rates of aggregate consumption C_t/C_{t-1} , stockholders consumption C_t^i/C_{t-1}^i , nonstockholders consumption C_t^n/C_{t-1}^n , stock returns R_t^s and the riskfree rate R_t^f in Guvenen’s model. Panel B reports properties of stochastic discount factors. The first row reports properties of the SDF for stockholders consumption. The remaining rows report SDF properties for total consumption and different preference parameters. The stochastic discount factors are of the CRRA form $M_t = \delta(C_t/C_{t-1})^{-\gamma}$. The first parameter in parenthesis is δ , the second one is γ . Panel C reports the annual pricing error Guvenen’s model. The preference parameters δ and γ are chosen to minimize the equally weighted sum of pricing errors for the stock returns R^s and the riskfree rate R^f . The first row labelled “SH” reports the pricing errors for stockholders consumption. The remaining rows labelled “AC” report pricing errors for aggregate consumption and different preference parameters. All statistics are quarterly.

Table 6: Limited Participation Habit Model with Habit SDF

σ_i/σ_c	$\rho(C_t^i/C_{t-1}^i, C_t/C_{t-1})$					
	-1.0	-0.5	-0.25	0.25	0.5	1.0
	δ_c					
1	0.85	0.73	0.32	0.54	0.85	0.96
2	0.88	0.75	0.56	0.62	0.90	0.97
5	0.84	0.68	0.36	0.72	0.99	1.02
	γ_c					
1	-1.18	-1.70	-2.43	5.23	2.50	0.83
2	-2.07	-3.41	-4.78	8.51	3.96	1.63
5	-4.72	-8.98	-16.54	19.04	9.29	4.36
	$E[M_t^{ch}R_t^s] - 1$ (in %)					
1	0.00	0.00	0.00	0.00	0.00	0.00
2	0.00	0.00	0.00	0.00	0.00	0.00
5	0.00	0.00	0.00	0.00	0.00	0.00
	$E[M_t^{ch}R_t^f] - 1$ (in %)					
1	0.00	0.00	0.00	0.00	0.00	0.00
2	0.00	0.00	0.00	0.00	0.00	0.00
5	0.00	0.00	0.00	0.00	0.00	0.00

Notes: This table reports preference parameters and pricing errors in Menzly, Santos and Veronesi's (2004) habit model. Consumption growth of stockholders is assumed to follow a random walk with a mean of 2% and standard deviation of 1%. All parameters are as in Menzly, Santos and Veronesi except α , which is set obtain the same average stock return as in Menzly-Santos-Veronesi. The preference parameters δ_c and γ_c are chosen to minimize the mean square pricing error $\min_{\delta, \gamma} [g(\delta, \gamma)' W g(\delta, \gamma)]$ where $g(\delta, \gamma) = E[M_t^{ch} \mathbf{R}_t - 1]$, $M_t^{ch} = \delta (\frac{C_t}{C_{t-1}} \frac{Y_{t-1}}{Y_t})^{-\gamma}$, $\mathbf{R} = [R^s, R^f]'$. C_t is aggregate consumption, Y_t is the inverse of the consumption surplus ratio computed from aggregate consumption, R^s is the return of equity, R^f is the riskfree rate, and $W = I$.

Table 7: Limited Participation Habit Model with CRRA SDF

σ_i/σ_c	$\rho(C_t^i/C_{t-1}^i, C_t/C_{t-1})$					
	-1.0	-0.5	-0.25	0.25	0.5	1.0
	δ_c					
1	0.55	0.29	0.07	4.41	2.38	1.56
2	0.55	0.28	0.06	4.54	2.44	1.59
5	0.52	0.25	NA	5.03	2.66	1.68
	γ_c					
1	-25.96	-52.40	-105.62	101.67	51.91	25.96
2	-26.72	-53.93	-108.87	104.56	53.44	26.72
5	-29.40	-59.33	NA	114.78	58.84	29.40
	$E[M_t^c R_t^s] - 1$ (in %)					
1	0.00	0.00	0.00	0.00	0.00	0.00
2	0.00	0.00	0.00	0.00	0.00	0.00
5	0.00	0.00	NA	0.00	0.00	0.00
	$E[M_t^c R_t^f] - 1$ (in %)					
1	0.00	0.00	0.00	0.00	0.00	0.00
2	0.00	0.00	0.00	0.00	0.00	0.00
5	0.00	0.00	NA	0.00	0.00	0.00

Notes: This table reports preference parameters and pricing errors in Menzly, Santos and Veronesi's (2004) habit model. Consumption growth of stockholders is assumed to follow a random walk with a mean of 2% and standard deviation of 1%. All parameters are as in Menzly, Santos and Veronesi except α , which is set obtain the same average stock return as in Menzly-Santos-Veronesi. The preference parameters δ_c and γ_c are chosen to minimize the mean square pricing error $\min_{\delta, \gamma} [g(\delta, \gamma)' W g(\delta, \gamma)]$ where $g(\delta, \gamma) = E[M_t^c \mathbf{R}_t - 1]$, $M_t^c = \delta (\frac{C_t}{C_{t-1}})^{-\gamma}$, $\mathbf{R} = [R^s, R^f]'$. C_t is aggregate consumption, R^s is the return of equity, R^f is the riskfree rate, and $W = I$. "NA" indicates that the numerical minimization did not converge.

Table 8: Limited Participation Habit Model with CRRA SDF and Larger Cross Section

σ_i/σ_c	$\rho(C_t^i/C_{t-1}^i, C_t/C_{t-1})$					
	-1.0	-0.5	-0.25	0.25	0.5	1.0
	δ_c					
1	0.99	0.97	0.93	1.02	1.01	1.01
2	0.99	0.97	0.97	1.02	1.01	1.01
5	0.99	0.97	0.97	1.02	1.01	1.01
	γ_c					
1	-29.47	-54.86	-116.62	123.39	64.28	30.41
2	-25.48	-63.03	-117.93	111.78	61.78	31.53
5	-37.39	-59.39	-133.63	135.24	64.45	32.47
	RMSE/ $E[\mathbf{R}_t]$ (in %)					
1	2.36	2.11	1.55	1.72	4.04	2.19
2	2.28	2.44	0.98	0.53	0.79	1.35
5	0.80	0.65	0.75	0.61	0.83	0.53

Notes: This table reports preference parameters and pricing errors in Menzly, Santos and Veronesi's (2004) habit model. Consumption growth of stockholders is assumed to follow a random walk with a mean of 2% and standard deviation of 1%. All parameters are as in Menzly, Santos and Veronesi except α , which is set obtain the same average stock return as in Menzly-Santos-Veronesi. The preference parameters δ_c and γ_c are chosen to minimize the mean square pricing error $\min_{\delta, \gamma} [g(\delta, \gamma)' W g(\delta, \gamma)]$ where $g(\delta, \gamma) = E[M_t^c \mathbf{R}_t - 1]$, $M_t^c = \delta (\frac{C_t}{C_{t-1}})^{-\gamma}$. C_t is aggregate consumption. \mathbf{R} includes the return of the market R^s , the riskfree rate R^f and the returns of six individual assets. RMSE/ $E[\mathbf{R}_t]$ gives the RMSE as a percentage of the cross sectional mean return of the eight assets. The weighting matrix W is the identity matrix.

Table 9: Limited Participation Habit Model with CRRA SDF and Time Aggregated Data

σ_i/σ_c	$\rho(C_t^i/C_{t-1}^i, C_t/C_{t-1})$					
	-1.0	-0.5	-0.25	0.25	0.5	1.0
	δ_c					
1	0.16	0.03	0.00	15.26	10.65	3.56
2	0.16	0.03	0.00	15.33	10.73	3.61
5	0.15	0.02	NA	15.55	11.09	3.82
	γ_c					
1	-77.04	-136.32	-270.70	257.24	180.54	77.16
2	-77.84	-138.31	-274.43	258.74	181.17	77.96
5	-80.96	-145.70	NA	264.13	184.56	81.08
	$E[M_t^c R_t^s] - 1$ (in %)					
1	0.00	0.00	0.00	0.00	0.00	0.00
2	0.00	0.00	0.01	0.00	0.00	0.00
5	0.00	0.00	NA	0.00	0.00	0.00
	$E[M_t^c R_t^f] - 1$ (in %)					
1	0.00	0.00	0.00	0.00	0.00	0.00
2	0.00	0.00	0.01	0.00	0.00	0.00
5	0.00	0.00	NA	0.00	0.00	0.00

Notes: This table reports preference parameters and pricing errors in Menzly, Santos and Veronesi's (2004) habit model. Consumption growth of stockholders is assumed to follow a random walk with a mean of 2% and standard deviation of 1%. All parameters are as in Menzly, Santos and Veronesi except α , which is set obtain the same average stock return as in Menzly-Santos-Veronesi. The preference parameters δ_c and γ_c are chosen to minimize the mean square pricing error $\min_{\delta, \gamma} [g(\delta, \gamma)' W g(\delta, \gamma)]$ where $g(\delta, \gamma) = E[M_t^c \mathbf{R}_t - 1]$, $M_t^c = \delta (\frac{C_t}{C_{t-1}})^{-\gamma}$, $\mathbf{R} = [R^s, R^f]'$. C_t is aggregate consumption, R^s is the return of equity, R^f is the riskfree rate, and $W = I$. "NA" indicates that the numerical minimization did not converge. The model is simulated on a quarterly frequency. The pricing errors are computed using the growth rate of annual consumption (the sum of four subsequent quarters) and compounded annual returns.

Table 10: Lim. Partic./Inc. Markets Pricing Errors for Stock Return and Risk-Free Rate: Hermite Densities

Cov($\Delta c, \Delta d$)=0.00017															
γ	δ	$\rho(\Delta c, \Delta c')$	$\rho(\Delta c', \Delta d)$	$\sigma(i)/\sigma(c)$	$\mu(\Delta c')/\mu(\Delta c)$	γ_c	δ_c	PrErrR(s)	PrErrR(f)	Sk[c]	Ku[c]	Sk[j]	Ku[j]	Sk[d]	Ku[d]
5	0.99	0.1	0.9	1	0.85	36.211	2.5613	4.95E-10	4.90E-10	4.0917	11.195	-0.0042337	3	0.036111	3.0009
5	0.99	0.1	0.9	1	1.5	36.217	2.3988	2.12E-10	2.10E-10	4.0899	11.181	-0.004228	3	0.036063	3.0009
5	0.99	0.1	0.9	2	0.85	71.495	6.0675	1.14E-09	1.12E-09	4.0952	11.207	0.0078509	3	0.04699	3.0015
5	0.99	0.1	0.9	2	1.5	71.53	5.6869	1.50E-09	1.49E-09	4.0934	11.193	0.0078403	3	0.046927	3.0015
5	0.99	0.1	0.9	4	0.85	129.08	14.235	9.75E-08	9.67E-08	4.1018	11.229	0.032021	3.0007	0.068751	3.0032
5	0.99	0.1	0.9	4	1.5	129.22	13.395	-9.01E-08	-8.50E-08	4.1	11.215	0.031977	3.0007	0.068658	3.0031

Cov($\Delta c, \Delta d$)=-0.00017															
γ	δ	$\rho(\Delta c, \Delta c')$	$\rho(\Delta c', \Delta d)$	$\sigma(i)/\sigma(c)$	$\mu(\Delta c')/\mu(\Delta c)$	γ_c	δ_c	PrErrR(s)	PrErrR(f)	Sk[c]	Ku[c]	Sk[j]	Ku[j]	Sk[d]	Ku[d]
5	0.99	0.1	0.9	1	0.85	-64.06	0.1052	-1.41E-14	-1.40E-14	4.1037	11.313	-0.0041596	3	-0.66437	3.2899
5	0.99	0.1	0.9	1	1.5	-64.05	0.0987	-1.27E-08	-1.27E-08	4.1034	11.303	-0.0041543	3	-0.66356	3.2891
5	0.99	0.1	0.9	2	0.85	-118.7	0.0121	-6.34E-14	-6.22E-14	4.1071	11.324	0.0077134	3	-0.65304	3.2802
5	0.99	0.1	0.9	2	1.5	-118.6	0.0113	-8.05E-11	-8.07E-11	4.1067	11.314	0.0077036	3	-0.65225	3.2795
5	0.99	0.1	0.9	4	0.85	-210.4	0.0002	-6.58E-13	-6.60E-13	4.1134	11.346	0.031459	3.0007	-0.63043	3.2614
5	0.99	0.1	0.9	4	1.5	-210.2	0.0001	-2.31E-13	-2.31E-13	4.1131	11.336	0.03142	3.0007	-0.62966	3.2607

Cov($\Delta c, \Delta d$)=0.00017															
γ	δ	$\rho(\Delta c, \Delta c')$	$\rho(\Delta c', \Delta d)$	$\sigma(i)/\sigma(c)$	$\mu(\Delta c')/\mu(\Delta c)$	γ_c	δ_c	PrErrR(s)	PrErrR(f)	Sk[c]	Ku[c]	Sk[j]	Ku[j]	Sk[d]	Ku[d]
5	0.99	0.1	0.9	1	0.85	35.488	1.7114	-7.93E-09	-8.00E-09	0.2175	3.0307	0.49692	2.4804	0.46153	2.2946
5	0.99	0.1	0.9	1	1.5	35.488	1.6032	-7.94E-09	-7.99E-09	0.2175	3.0307	0.49691	2.4804	0.46153	2.2946
5	0.99	0.1	0.9	2	0.85	70.978	2.7445	9.82E-09	9.64E-09	0.2175	3.0307	0.49692	2.4804	0.46154	2.2946
5	0.99	0.1	0.9	2	1.5	70.978	2.571	9.82E-09	9.64E-09	0.2175	3.0307	0.49691	2.4804	0.46153	2.2946
5	0.99	0.1	0.9	4	0.85	141.96	4.3612	2.26E-07	2.25E-07	0.2175	3.0307	0.49692	2.4804	0.46154	2.2946
5	0.99	0.1	0.9	4	1.5	141.96	4.0855	2.26E-07	2.25E-07	0.2175	3.0307	0.49692	2.4804	0.46153	2.2946

Cov($\Delta c, \Delta d$)=-0.00017

γ	δ	$\rho(\Delta c, \Delta c^i)$	$\rho(\Delta c^i, \Delta d)$	$\sigma(i)/\sigma(c)$	$\mu(\Delta c^i)/\mu(\Delta c)$	γ_c	δ_c	PrErrR(s)	PrErrR(f)	Sk[c]	Ku[c]	Sk[i]	Ku[i]	Sk[d]	Ku[d]
5	0.99	0.1	0.9	1	0.85	-35.46	0.4115	-4.87E-08	-4.88E-08	-0.218	3.0308	0.49691	2.4804	0.46152	2.2946
5	0.99	0.1	0.9	1	1.5	-35.46	0.3855	-4.88E-08	-4.88E-08	-0.218	3.0308	0.4969	2.4804	0.46151	2.2946
5	0.99	0.1	0.9	2	0.85	-70.92	0.1587	4.66E-15	4.66E-15	-0.218	3.0308	0.49691	2.4804	0.46152	2.2946
5	0.99	0.1	0.9	2	1.5	-70.92	0.1487	4.22E-15	5.33E-15	-0.218	3.0308	0.4969	2.4804	0.46151	2.2946
5	0.99	0.1	0.9	4	0.85	-141.8	0.0146	1.21E-13	1.22E-13	-0.218	3.0308	0.49692	2.4805	0.46153	2.2946
5	0.99	0.1	0.9	4	1.5	-141.8	0.0137	1.15E-13	1.18E-13	-0.218	3.0308	0.49691	2.4804	0.46152	2.2946

Cov($\Delta c, \Delta d$)=0.00017

γ	δ	$\rho(\Delta c, \Delta c^i)$	$\rho(\Delta c^i, \Delta d)$	$\sigma(i)/\sigma(c)$	$\mu(\Delta c^i)/\mu(\Delta c)$	γ_c	δ_c	PrErrR(s)	PrErrR(f)	Sk[c]	Ku[c]	Sk[i]	Ku[i]	Sk[d]	Ku[d]
5	0.99	0.1	0.9	1	0.85	34.614	1.6892	-3.03E-10	-2.94E-10	0.2389	3.0373	0.54147	2.5713	0.50496	2.39
5	0.99	0.1	0.9	1	1.5	34.619	1.5825	-3.03E-10	-2.96E-10	0.2425	3.0385	0.55543	2.6008	0.51853	2.4204
5	0.99	0.1	0.9	2	0.85	69.241	2.6948	1.24E-11	1.20E-11	0.2377	3.0369	0.53729	2.5614	0.50128	2.3818
5	0.99	0.1	0.9	2	1.5	69.249	2.5247	1.26E-11	1.24E-11	0.2413	3.0381	0.55113	2.5905	0.51472	2.4118
5	0.99	0.1	0.9	4	0.85	138.64	4.3376	-2.37E-09	-1.87E-09	0.2353	3.0361	0.52911	2.542	0.49409	2.3659
5	0.99	0.1	0.9	4	1.5	138.64	4.0636	-2.40E-09	-1.89E-09	0.2389	3.0373	0.54269	2.5705	0.50727	2.3952

Cov($\Delta c, \Delta d$)=-0.00017

γ	δ	$\rho(\Delta c, \Delta c^i)$	$\rho(\Delta c^i, \Delta d)$	$\sigma(i)/\sigma(c)$	$\mu(\Delta c^i)/\mu(\Delta c)$	γ_c	δ_c	PrErrR(s)	PrErrR(f)	Sk[c]	Ku[c]	Sk[i]	Ku[i]	Sk[d]	Ku[d]
5	0.99	0.1	0.9	1	0.85	-36.69	0.3989	-2.05E-11	-2.08E-11	-0.218	3.0311	0.54207	2.5712	0.506	2.3923
5	0.99	0.1	0.9	1	1.5	-36.68	0.3737	-1.81E-11	-1.83E-11	-0.221	3.0321	0.55607	2.6007	0.51961	2.4228
5	0.99	0.1	0.9	2	0.85	-73.97	0.1449	-3.45E-10	-3.50E-10	-0.215	3.0303	0.53788	2.5613	0.50231	2.3841
5	0.99	0.1	0.9	2	1.5	-73.95	0.1358	-3.37E-10	-3.42E-10	-0.218	3.0313	0.55174	2.5905	0.51578	2.4142
5	0.99	0.1	0.9	4	0.85	-150.4	0.0104	0	1.33E-15	-0.209	3.0287	0.52967	2.5418	0.49508	2.3681
5	0.99	0.1	0.9	4	1.5	-150.4	0.0098	6.66E-16	4.44E-16	-0.212	3.0296	0.54328	2.5704	0.50829	2.3975

Notes: This table reports output on the pricing error associated with erroneously using aggregate consumption in place of asset-holder consumption, for a range of parameter values and joint distributions. γ^i is the presumed value of asset-holder risk-aversion; δ^i is the presumed value of the asset-holder's subjective discount rate; $\rho(\Delta c, \Delta c^i)$ denotes the correlation between aggregate consumption growth and asset-holder consumption growth in the leading normal; $\rho(\Delta c^i, \Delta d)$ denotes the correlation between asset-holder consumption growth and dividend growth in the leading normal; $\sigma(\Delta c^i)/\sigma(\Delta c)$ denotes the standard deviation of asset-holder consumption growth divided by the standard deviation of aggregate consumption growth in the leading normal; $\mu(\Delta c^i)/\mu(\Delta c)$ denotes the mean of asset-holder consumption growth divided by the mean of aggregate consumption growth in the leading normal; γ^c and δ^c are the values of γ and δ that minimize the pricing errors using aggregate consumption; PrErrR(s) is the pricing error for the Euler equation associated with the stock return; PrErrR(f) is the pricing error of the Euler equation associated with the risk-free rate, and Sk[], Ku[] refer to the skewness and kurtosis of aggregate consumption (c), asset-holder consumption (i), and dividends (d).

Table 11: Lim. Partic./Inc. Markets Pricing Errors in a Larger Cross-Section: Hermite Densities

Distribution	γ	δ	$\rho(\Delta c, \Delta c^i)$	$\sigma(\Delta c^i)/\sigma(\Delta c)$	$\mu(\Delta c^i)/\mu(\Delta c)$	$\text{Cov}(\Delta c, \Delta d)$	γ_c	δ_c	Max RMSE	Avg RMSE
J. Log N.	5	0.99	0.13	2	1.5	0.00017	78.24	2.77	0.02%	0.02%
Non-Normal	5	0.99	0.13	2	1.5	0.00017	6.83	1.09	10.10%	0.25%
J. Log N.	5	0.99	0.13	2	1.5	-0.00017	-78.24	0.12	0.02%	0.02%
Non-Normal	5	0.99	0.13	2	1.5	-0.00017	-82.8	0.209	0.58%	0.03%

Notes: This table reports the average pricing errors for models with 8 asset returns. The column labeled "Distribution" denotes whether the joint distribution of Δc , Δc^i and dividend growth for each of the 8 assets is modeled as lognormal or not. "J. Log N." reports results for the jointly lognormal case; "Non-Normal" reports the results for cases in which a perturbation from the lognormal was used to describe the joint distribution of aggregate consumption, asset-holder consumption, and the 8 asset returns. The numbers in the column labeled "Max RMSE" give the square root of the average squared pricing error, as a fraction of the cross-sectional average mean return, that is the *maximum* over all Non-Normal perturbations (over 100) considered. The numbers in the column labeled "Avg RMSE" give the square root of the average squared pricing error, as a fraction of the cross-sectional average mean return, that is the *average* of over Non-Normal perturbations (over 100) considered. γ^i is the presumed value of asset-holder risk-aversion; δ^i is the presumed value of the asset-holder's subjective discount rate; $\rho(\Delta c, \Delta c^i)$ denotes the correlation between aggregate consumption growth and asset-holder consumption growth in the leading normal; $\rho(\Delta c^i, \Delta d)$ denotes the correlation between asset-holder consumption growth and dividend growth in the leading normal; $\sigma(\Delta c^i)/\sigma(\Delta c)$ denotes the standard deviation of asset-holder consumption growth relative to the standard deviation of aggregate consumption growth in the leading normal; $\mu(\Delta c^i)/\mu(\Delta c)$ denotes the mean of asset-holder consumption growth divided by the mean of aggregate consumption growth in the leading normal. γ^c and δ^c are the values of γ and δ that minimize the equally weighted sum of squared pricing errors when aggregate consumption is used in place of stockholder consumption, for the hermite distribution that delivers the *maximum* RMSE, as a percentage of the cross-sectional mean return. For the jointly lognormal case, the average is the maximum since there is only one distribution to average over.