Euler Equation Errors[*]

Martin Lettau

New York University, CEPR, NBER

Sydney C. Ludvigson

New York University and NBER

Comments Welcome

First draft: September 1, 2004

This draft: August 9, 2005

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 cross-sections of asset returns. Here we ask whether calibrated leading asset pricing models—specifically developed to address empirical puzzles associated with the standard paradigm—explain the mispricing of the standard consumption-based model when evaluated on cross-sections of asset returns. 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, such a kernel will not rationalize the magnitude of the pricing errors generated by the standard model, particularly when the curvature of utility is high. Second, we show that leading asset pricing models also do not explain the significant mispricing of the standard paradigm for plausibly calibrated sets of asset returns, 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 one example of a limited participation/incomplete markets model capable of explaining larger pricing errors for the standard model; but we also find many examples of such models, in which the consumption of marginal assetholders behaves quite differently from per capita aggregate consumption, that do not explain the large Euler equation errors of the standard representative agent model.

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 cross-sections of asset returns. Such Euler equation errors, or pricing errors (terms we use interchangeably), are especially large for cross-sections that include a broad stock market index return, a short term Treasury bill rate, and the size and book-market sorted portfolio returns emphasized by Fama and French (1992,1993). The large unconditional pricing equation errors of the standard model have been stressed elsewhere as an indication of the model's empirical difficulties, e.g., Mankiw and Shapiro (1986), Breeden, Gibbons, and Litzenberger (1989), Campbell (1996), Cochrane (1996), and Lettau and Ludvigson (2001). We present further evidence on the size of these errors here.

The *standard* model, as we define it, 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

$$E\left[M_{t+1}R_{t+1}\right] = 1, \qquad M_{t+1} = \delta\left(C_{t+1}/C_t\right)^{-\gamma}, \tag{1}$$

where the first equality is the Euler equation, R_{t+1} denotes the gross return on any tradable asset, C_{t+1} is per capita aggregate consumption, γ is the coefficient of relative risk-aversion and δ is a subjective time-discount factor. For any model, the Euler equation error of a tradable asset is the difference between $E\left[M_{t+1}R_{t+1}\right]$ and unity.

In much of theoretical asset pricing, Euler equations give the equilibrium prices of tradable assets, consequently the empirical errors in this equation are a fundamental measure of how well any model explains asset returns. It is perhaps surprising, then, that little research has been devoted to investigating whether leading asset pricing models can help us understand the significant mispricing of the standard model when confronted with cross-sections of stock returns. After all, such models were developed with the express purpose of explaining the empirical limitations of the standard model. Instead, theoretical research has proceeded by focusing on well known "puzzles" generated by the standard model, for example, the equity premium puzzle, the risk-free rate puzzle, and the time-series predictability of excess stock market returns.

In this paper, we argue that large empirical Euler equation errors constitute a puzzle for the standard consumption-based asset pricing model that is at least as damning as these other, more well known, conundrums. We employ the empirical facts on Euler equation errors to evaluate some leading asset pricing models that were specifically developed to

address puzzles associated with the standard paradigm (1). This is of interest because 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 particular, 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.

We find that this is not always the case. Often, in leading asset pricing models, parameters of a standard representative agent "pricing kernel" based on (1) can be found that imply the standard model has virtually identical unconditional pricing implications as the true model that prices assets correctly. Thus, an econometrician who observed data generated from any 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 (1) to historical data.[1]

We note that the literature has already demonstrated 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 (Constantinides and Duffie (1996)). Such existence proofs, important in their own right, are not the focus of this paper. Instead, we ask whether particular calibrated economies of leading asset pricing models are quantitatively capable of matching the large pricing equation errors generated by the standard consumption-based model when fitted to historical data. This is important because it remains unclear whether fully specified models built on primitives of tastes, technology, and underlying shocks, and calibrated to accord with the data in plausible ways, can in practice generate the joint behavior of aggregate consumption and asset returns that we observe in the data.

Our analysis uses simulated data from several leading asset pricing models: 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 risk 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, and each has proven remarkably successful in explaining a range of

---

[1]Note that these findings are not a statement about the "power" of empirical Euler equation errors to reject alternative models, just as evidence that most business cycle models fail to explain a large equity premium and a low and stable risk-free rate is not a statement about the power of key asset pricing moments to reject alternative business cycle models. The interesting finding is not that the alternative models we study have similar properties along this dimension, but rather that they all fail in the same way to explain the significant mispricing of the standard model.

asset pricing phenomena that the standard model is unable to explain.

Our focus on Euler equations is intentional, since they represent the set of theoretical restrictions from which all asset pricing implications follow. Kocherlakota (1996) emphasizes the importance of Euler equation errors for understanding the central empirical puzzles of the standard consumption-based model, which he illustrates using annual data on aggregate consumption and asset returns. 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 stress here that the quarterly pricing errors for the *unconditional* Euler equations associated with cross-sections of asset returns are large when fitting aggregate data to (1), even when the parameters δ and γ are left unrestricted and chosen to minimize those errors. Such Euler equation errors 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.

Our main findings are as follows:

First, we consider the case in which consumption in (1) is mismeasured, perhaps because per capita aggregate consumption is a poor measure of individual assetholder consumption, or the consumption of stockholders as an aggregate. We show that if the true pricing kernel based on assetholder consumption is jointly lognormally distributed with aggregate consumption and returns, then estimation of (1) using per capita aggregate consumption produces biased estimates of the assetholder's subjective discount factor and risk aversion parameters, but does not rationalize the magnitude of the pricing errors generated by the standard model, particularly when γ is large.

Second, we use simulated data from each of the leading asset pricing models mentioned above to study the extent to which these models explain the mispricing of the standard model. We show that some of these models *can* explain why an econometrician obtains implausibly high estimates of δ and γ when freely fitting aggregate data to (1). But, none can explain the large unconditional Euler equation errors associated with such estimates for plausibly calibrated sets of asset returns.[2] Indeed, the asset pricing models we consider

---

[2]Campbell and Cochrane (2000) evaluate the pricing errors of the standard consumption-based model implied by the habit model of Campbell and Cochrane (1999), by looking at the pricing errors for the *most mispriced* portfolio. Their results suggest that there is scope for mispricing, but do not imply significant mispricing for the sets of stock portfolios we calibrate our models to match. Our approach differs from theirs in that we do not analyze the most mispriced portfolio (which can look nothing like the stock portfolios

counterfactually imply that values of δ and γ can be found for which (1) satisfies the unconditional Euler equation restrictions just as well as the true pricing kernel, implying that the standard model generates negligible pricing errors for cross-sections of asset returns.

Third, in contrast to the above results, we provide one example of an incomplete markets/limited participation model that can rationalize larger pricing errors for the standard model, as long as the joint distribution of aggregate consumption, individual assetholder consumption, and stock returns takes a particular non-normal form. But we also find—within the class of distributions we consider—that many models with non-normal distribution specifications will not explain mispricing of the standard model, since in many cases the use of (1) to price assets merely distorts the estimated preference parameters but not the pricing errors.

We emphasize that this paper is not a criticism of existing asset pricing theory. Instead, we seek a diagnostic for understanding the directions in which future research may be fruitfully applied by providing a different perspective on whether leading paradigms fully rationalize the joint behavior of asset prices and aggregate quantities that is central to the empirical failure of the standard model. 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. Our point is that many models written down today appear inconsistent with these restrictions and do not explain the mispricing of the standard consumption-based model.

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, beginning 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 the four leading asset pricing models mentioned above are capable of explaining the empirical facts. Our main findings are shown to be robust to time-aggregation of aggregate consumption data, to the introduction of limited participation in the representative agent models, and to the use of small samples to compute pricing errors. Finally, we explore the pricing implications of a number of simple incomplete markets/limited participation models in which assetholder consumption

_____

observed in historical data), but instead generate specific cross-sections of traded assets in the models to match the properties of cross-sections in the data by directly relying on the models' own baseline calibrations of asset returns, or by employing calibrations which deliver spreads in risk-premia commensurate with those in our historical data set.

is permitted to behave quite differently from per capita aggregate consumption. Section 4 concludes.

2 Euler Equation Errors: Empirical Facts

The standard consumption based model, as defined above, assumes a representative-consumer with constant relative risk aversion (CRRA) preferences over consumption given by

$$U = E\left\{\sum_{t=0}^{\infty} \delta^t \frac{C_t^{1-\gamma} - 1}{1 - \gamma}\right\}, \qquad \gamma > 0. \tag{2}$$

At each date, agents maximize (2) subject to an accumulation equation for wealth. Agents have unrestricted access to financial markets and face no borrowing or short-sales constraints. The asset pricing model comes from the first-order conditions for optimal consumption choice, which imply that for any traded asset indexed by j, with a gross return at time $t+1$ of R_{t+1}^j, the following Euler equation holds:

$$E_t\left[\delta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} R_{t+1}^j\right] = 1. \tag{3}$$

Here E_t is the conditional expectation operator, conditional on time t information. The marginal rate of substitution in consumption, $M_{t+1} \equiv \delta\left(C_{t+1}/C_t\right)^{-\gamma}$, is the stochastic discount factor, or pricing kernel in this model. By the law of iterated expectations, equation (3) also implies a corresponding unconditional Euler equation taking the form (1), which we focus on from here on.

We focus our attention on the unconditional Euler equation errors for cross-sections of asset returns that include a broad stock market index return (measured as the CRSP value-weighted price index return and denoted R_t^s), a short term Treasury bill rate (measured as the three-month Treasury bill rate and denoted R_t^f), and six size and book-market sorted portfolio returns available from Kenneth French's Dartmouth web site. These returns are value-weighted portfolio returns of common stock sorted into two size (market equity) quantiles and three book value-market value quantiles. We use equity returns on 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. These returns are denoted as a vector $\mathbf{R}_t^{FF} \equiv (R_t^1, ... R_t^6)'$. We analyze the pricing errors for the eight assets $R_t^s, R_t^f, \mathbf{R}_t^{FF}$ as a group, as well as for the set of two assets comprised of only R_t^s and R_t^f. The latter is of interest because the standard model's inability to explain properties of these two returns

5

has been central to the development of a consensus that the model is flawed. In addition, almost all asset pricing models seek to match the empirical properties of these two returns, whereas fewer generate implications for larger cross-sections of securities.

There are two ways to present the pricing errors implied by the standard consumption-based model. One is to focus on the Euler equations of raw returns:

$$E\left[\delta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}R_{t+1}^j\right] - 1 = 0 \qquad j = s, f, 1, ..., 6. \tag{4}$$

Another is to focus on the Euler equation errors for excess returns:

$$E\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\left(R_{t+1}^j - R_{t+1}^f\right)\right] = 0 \qquad j = s, 1, ..., 6. \tag{5}$$

For both Euler equations above, we refer to the difference between an estimate of the left-hand-side and zero as the *unconditional Euler equation error*, or alternatively the *pricing error*, for the jth asset return. 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.

Regardless of whether the Euler equations are stated in terms of excess or raw returns, we choose the parameters δ and γ to minimize a weighted sum of squared pricing errors, an application of Generalized Method of Moments (GMM, Hansen (1982)):

$$\min_{\delta,\gamma}\ g_T\left(\gamma,\delta\right) \equiv \mathbf{w}_T'\left(\gamma,\delta\right)\mathbf{W}\mathbf{w}_T\left(\gamma,\delta\right), \tag{6}$$

where $\mathbf{W}$ is the identity matrix and $\mathbf{w}_T\left(\gamma,\delta\right)$ is the vector of average pricing errors for each asset, with jth element $w_{jT}(\gamma,\delta)$ given either by

$$w_{jT}(\gamma) = w_{jT}\left(\gamma\right) = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\left(R_{t+1}^j - R_{t+1}^f\right), \qquad j = s, 1, ..., 6.$$

in the case of excess returns, or

$$w_{jT}(\gamma,\delta) = \frac{1}{T}\sum_{t=1}^{T}\delta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}R_{t+1}^j - 1, \qquad j = s, f, 1, ..., 6$$

in the case of raw returns. Let $\widehat{\delta}$ and $\widehat{\gamma}$ denote the $\arg\min g_T\left(\gamma,\delta\right)$.

We use the identity weighting matrix because these it preserves the structure of the original test assets, which are based on economically interesting characteristics and deliver a wide spread in cross-sectional average returns. Use of alternate matrixes that re-weight the Euler equations amount to minimizing the pricing errors for re-weighted portfolios of the original test assets and destroy this structure. It should be noted, however, that other weighting matrixes such as the optimal weighting matrix of Hansen (1982) and the second

moment matrix of Hansen and Jagannathan (1997) produce similar results although they are not reported in what follows.

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.[3] Returns are deflated by the implicit price deflator corresponding to the measure of consumption C_t. The data span the period from the fourth quarter of 1951 to the fourth quarter of 2002. A detailed description of the data is provided in the Appendix.

Figure 1 displays the square root of the average squared Euler equation errors (RMSE) for the excess returns in (5) over a range of values of γ. The solid line plots the case where the single excess return $R_{t+1}^s - R_{t+1}^f$ is priced; the dotted line plots the case for the seven returns returns $R_{t+1}^s - R_{t+1}^f$ and $\mathbf{R}_t^{FF} - R_{t+1}^f$. To give a sense of how large pricing errors are relative to the returns being priced, we plot RMSE/RMSR, where RMSR is the square root of the cross-sectional average of the squared mean returns of the assets under consideration.

Two aspects of Figure 1 warrant emphasis. First, notice that in the case of the single excess return on the aggregate stock market, $R_{t+1}^s - R_{t+1}^f$ (solid line), the RMSE is itself just the pricing error (5), where this error is computed as the sample mean of the expression in square brackets in (5), scaled by the value of δ that minimizes an equally weighted average of Euler equation errors for R_t^s and R_t^f. The solid line shows that the pricing error (5) for the excess return on the aggregate stock market cannot be driven to zero, or indeed even to a small number, for any value of γ. The lowest pricing error is 5.2% per annum, which occurs at $\gamma = 117$. The figure displays this error as a fraction of the average excess stock market return, and is shown to be almost 60 percent of the average annual CRSP excess return. 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.[4]

Second, the dashed line in Figure 1 shows that the root mean-squared pricing errors for the seven asset case $R_{t+1}^s - R_{t+1}^f, \mathbf{R}_t^{FF} - R_{t+1}^f$ is also large. As a fraction of the square root of the average squared excess returns being priced, the minimum RMSE is about 60%, about the same as that for the single excess return $R_{t+1}^s - R_{t+1}^f$, and this occurs at $\gamma = 118$. At other values of γ the RMSE rises precipitously, just as it does for the single asset case. Therefore, the degree of mispricing in the standard model is about the same regardless of whether we consider the single excess return on the market or a larger cross-section of excess stock market returns.[5]

---

[3]We exclude shoes and clothing expenditure from this series since they are partly durable and should not be included in a measure of the service flow of consumption.

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

[5]In computing the pricing errors above, we use the standard timing convention that end-of-period returns

Next we report the Euler equation errors in (4) for raw returns. Table 1 shows that when δ and γ are chosen to minimize (6) for R^s_{t+1} and R^f_{t+1} alone, the RMSE is 2.7% per annum, a magnitude that is 48% of the square root of the average squared returns on these two assets. Since there are just two moments in this case, this again means that there are no values of δ and γ that set the two pricing errors to zero. When δ and γ are chosen to minimize (6) for the eight asset returns $R^s_{t+1}, R^f_{t+1}, \mathbf{R}^{FF}_t$, the RMSE is 3.05% per annum, a magnitude that is 33% of the square root of the average squared returns on the eight assets. Notice that the estimates $\widehat{\delta}$ and $\widehat{\gamma}$ (which are left unrestricted) are close to 1.4 and 90, respectively, regardless of which set of test assets are used. The final two columns of Table 1 report the results of statistical tests of the model, discussed below.

Why are the pricing errors so large? The lower panel of Table 1 provides a partial answer: a significant part of the unconditional Euler equation errors generated by the standard model are associated with recessions, periods in which per capita aggregate consumption growth is steeply negative. For example, if we remove the data points associated with the smallest six observations on consumption growth, the RMSE is 0.73% per annum or 13% of the root mean squared returns for R^s_{t+1} and R^f_{t+1}, and 1.94% per annum or 21 percent of the root mean squared returns on the eight asset returns $R^s_{t+1}, R^f_{t+1}, \mathbf{R}^{FF}_t$. 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. This result echoes the findings in Ferson and Merrick (1987) who report less evidence against the standard consumption-based model in non-recession periods. 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+1} and returns) contributes to large pricing errors.[6] One can also reduce the pricing errors by using annual

---

dated in quarter t should be paired with consumption growth measured from $t - 1$ to t. If, instead, returns at t are paired with consumption growth from t to $t + 1$, a value for γ can be found that sets the pricing error to zero for the single excess return $R^s - R^f$. By contrast, the choice of timing convention has very little affect on the RMSE for the set of seven excess returns $R^s - R^f, \mathbf{R}^{FF}_t - R^f$. We use the former timing convention as it is standard empirical practice in estimation of Euler equations. We stress, however, that the timing convention itself is not important for the comparisons with theoretical models that follow, since those models always produce zero pricing errors *regardless* of which timing convention is used.

[6]Eliminating the recession periods, however, results in preference parameter estimates that are even more extreme than they are in the full sample; for example $\gamma_c > 300$. Therefore, if one's criterion for success is reasonable preference parameter estimates, then the standard model does worse when the recession periods are removed than when they are included. If γ is restricted to be less than 100 in the sample without recessions, the pricing errors move up considerably. For example, in the two asset case the RMSE moves up to 1.94% from 0.73%.

returns and year-over-year consumption growth.[7] This procedure averages out the worst quarters for consumption growth instead of removing them. Either procedure eliminates a substantial proportion of the cyclical variation in consumption. For example, 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. This explains why Kocherlakota (1996), who focuses on annual data, is able to locate parameter values for δ and γ that exactly satisfy the Euler equations (4) for a stock return and Treasury-bill rate.

Of course, these quarterly recession 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.

Although not reported above, we note that the pricing error of the Euler equation associated with the CRSP stock market return is always positive, implying a positive "alpha" in the expected return-beta representation of the model.[8] 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. 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.

2.1 Sampling Error and Tests for Joint Normality

We can use GMM distribution theory to ask whether the estimated pricing errors $\mathbf{w}_T (\gamma, \delta)$ are jointly different from zero, that is larger than what would be implied by sampling error alone. When there are more moments than parameters to be estimated, such an assessment can be interpreted as a test of overidentifying restrictions. The last two columns of Table 1 report $p-$values from chi-squared tests of the model's overidentifying restrictions for estimation of the eight Euler equations in (4). Although the results presented so far have used the identity weighting matrix, the last column in Table 1 presents the $p-$values from the same statistical test using an estimate of the optimal GMM weighting matrix (Hansen (1982)). The results from either weighting matrix are the same: we may strongly reject the hypothesis that the Euler equation errors are jointly statistically indistinguishable from zero;

---

[7] Jagannathan and Wang (2004) study the ability of a linearized version of the standard model to explain a large cross-section of asset returns using forth 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.

[8] The alpha in the expected return-beta representation is equal to the pricing error, scaled by $1/E[M_t]$; see Cochrane (2005) for an exposition.

the $p-$values for this test are less than 0.0001.

For the two-asset case, the model is just-identified, so the overidentifying tests above are not applicable. But note that the expectation in (5) is estimated using the sample means of

$$e_{t+1} = \delta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \left(R^s_{t+1} - R^f_{t+1} \right),$$

which are excess returns discounted by the pricing kernel $\delta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma}$. Fixing δ and γ, it is possible to compute the sampling variation in the sample mean of e_{t+1}, given as $\sigma^2 = \sigma^2_e/T$, where σ_e is the sample standard deviation of e_{t+1} and T is the sample size.[9] Not surprisingly, the sampling error of the mean of e_{t+1} is quite large when evaluated at the optimal values of $\delta = 1.4$ and $\gamma = 117$: a confidence interval formed by plus and minus two standard errors is $(-0.55\%, 11\%)$ as a percent per annum. This large range, which includes zero, arises partly for the same reason that it is difficult to estimate the equity premium accurately: excess returns are highly volatile. But it is also because the data force a very high value for γ in an attempt to fit the equity premium. Such a high value of γ generates extreme volatility in the pricing kernel, making discounted returns even harder to estimate precisely than nondiscounted returns. Unless one views $\gamma = 117$ as plausible, however, such wide standard error bands for mean discounted returns merely provide further evidence of the model's empirical limitations, which even at $\gamma = 117$ leaves a pricing error that is more than half of the average annual stock return. If instead we restrict the value of risk aversion to lie in the range $0 \leq \gamma \leq 89$, the pricing errors are always statistically different from zero at the five percent level of significance. In short, when γ is as high as 117, the sample mean of e_{t+1} is statistically insignificant not because the pricing errors are small–indeed they are economically large–but rather because discounted returns are so extremely noisy when $\gamma = 117$.

For the case of raw returns and only two assets R^s_{t+1} and R^f_{t+1}, we 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?[10] Models that postulate joint lognormality for consumption and asset returns

---

[9]We also calculated standard errors for the mean of e_{t+1} using a nonparameteric correction for serial correlation. Since e_{t+1} is close to serially uncorrelated, this correction has little affect on the error bands.

[10]For the case of raw returns and only two assets R^s_{t+1} and R^f_{t+1}, we have an exactly identified GMM system, so sampling error could in principle be assessed 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.

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 for two asset returns. To address the question just raised, we suppose the data were generated by the standard CRRA representative agent model, with returns and consumption jointly lognormally distributed, and ask how likely is it that we would find results like those reported in Table 1, in a sample of the size we have.

Consider 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. 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 (204), 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 (9.5×10^{-11}, 7.0×10^{-9}) for the risky return and (1.3×10^{-10}, 6.5×10^{-9}) for the risk-free return. These findings suggest that it is extremely unlikely that we would find results like those reported in Table 1, in a sample of the size we have, if this simple version of the standard CRRA representative agent model, where consumption and returns are jointly lognormally distributed, were true.

Given these results, it is natural to assess 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 the data based on multivariate skewness and kurtosis for the vector $\mathbf{Y}_t \equiv \left[\log \left(C_{t+1}/C_t \right), \log \left(R^s_{t+1} \right), \log \left(R^f_t \right) \right]'$. We also perform joint normality tests for the larger set of variables

$$\mathbf{X}_t \equiv \left[\log \left(C_{t+1}/C_t \right), \log \left(R^s_{t+1} \right), \log \left(R^f_t \right), \log \left(R^1_t \right), ..., \log \left(R^6_t \right) \right].$$

Normality tests for the larger cross-section will help inform the results in the next section in which models that assume joint lognormality are studied.[11]

Statistical tests based on multivariate skewness and kurtosis provide strong evidence against joint normality. For $\mathbf{Y}_t$ multivariate skewness is estimated to be 1.54 and multivariate excess kurtosis is 4.64, with $p-$values for the null hypothesis that these statistics are equal to

---

[11]Multivariate skewness and kurtosis statistics are computed 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

those of a multivariate normal distribution less than 0.0001. Similarly for $\mathbf{X}_t$, multivariate skewness is 4.65 and multivariate kurtosis is 35.93, and the statistical rejections of normality are even stronger. The same conclusion arises from examining quantile-quantile plots (QQ plots) for the vector time-series $\mathbf{Y}_t$ and $\mathbf{X}_t$, given in Figure 3. This figure plots the sample quantiles for the data against those that would arise under the null of joint lognormality, along with pointwise standard errors bands.[12] The QQ plots show substantial departures from normality: a large number of quantiles lie far outside the standard error bands for joint normality.

3   Euler Equation Errors: The Theories

How capable are asset pricing theories of explaining the large pricing errors of the standard model? In this section, we address this question by considering a number of distinct asset pricing models. We begin with a simple model of limited participation/incomplete markets model in which the true pricing kernel based on assetholder consumption is jointly lognormally distributed with aggregate consumption and returns. Although the empirical results reported above suggest that any model that implies aggregate consumption and returns are jointly lognormally distributed will be unable to match the data, studying a lognormal model is instructive for considering how the use of a mismeasured pricing kernel (for example because per capita aggregate consumption is used in place of stockholder consumption) might distort parameters compared to pricing errors. Next we evaluate the Euler equation errors generated by leading asset pricing models in which the log pricing kernel and returns are not generally lognormally distributed. As mentioned, these include the external habit-formation models of Campbell and Cochrane (1999) and Menzly, Santos, and Veronesi (2004), the long-run risk model of Bansal and Yaron (2004), and the limited participation model of Guvenen (2003). Finally, we present a number of additional results for simple lim-

---

skewness S and (excess) kurtosis K and asymptotic distributions are given by

$$
\begin{aligned}
S &= \left(\frac{1}{T^2} \sum_{t=1}^{T} \sum_{s=1}^{T} g_{ts}^3 \right)^{1/2} & \frac{TS^2}{6} &\sim \chi^2_{p(p+1)(p+2)/6} \\
K &= \frac{1}{T} \sum_{t=1}^{T} g_{tt}^2 - p(p+2) & \frac{\sqrt{T}K}{\sqrt{8p(p+2)}} &\sim N(0,1),
\end{aligned}
$$

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.

[12]Pointwise standard error bands are computed by simulating from the multivariate normal distribution with length equal to the size of our data set.

ited participation/incomplete markets models in which assetholder consumption, aggregate consumption and asset returns are not jointly lognormally distributed.

3.1 A Limited Participation/Incomplete Markets Model With Joint Lognormality

We investigate the affect on parameter estimates and pricing errors of estimating (1) on aggregate consumption data when the return data were generated from a model with limited stock market participation or incomplete markets. For this purpose, a model of limited stock market participation is isomorphic to that of incomplete markets since what matters is the common implication that the consumption of the marginal assetholder may behave differently from per capita aggregate consumption.[13] Thus, 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.

As a benchmark case in this section, we assume aggregate consumption, stockholder or individual consumption, and asset returns are jointly lognormally distributed. Later we consider asset pricing models in which the joint distribution is permitted to deviate from lognormality. For the rest of the paper, we use lowercase letters to denote log variables, e.g., $\Delta c_{t+1} \equiv \log\left(C_{t+1}/C_t\right).$

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

$$M_{t+1}^i \equiv \delta \left(\frac{C_{t+1}^i}{C_t^i} \right)^{-\gamma}, \tag{7}$$

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+1}^i correctly prices any traded asset return, implying that

$$E\left[M_{t+1}^i R_{t+1}^j\right] = 1, \qquad j = 1, ..., N \tag{8}$$

for N asset returns.

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

---

[13]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. Similarly, with incomplete consumption insurance the set of Euler equations of *household* consumption imply that any household's marginal rate of substitution is a valid stochastic discount factor.

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)). It functions as the stochastic discount factor in this model. The risk-free rate is defined as a one-period riskless bond, $R_{t+1}^f = 1/E_t\left[M_{t+1}^i\right]$.

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

$$M_{t+1}^c \equiv \delta_c \left(\frac{C_{t+1}}{C_t}\right)^{-\gamma_c}. \tag{9}$$

For any asset return indexed by j, the pricing error associated with the true MRS, M_{t+1}^i, is by construction zero, but the pricing error associated with the erroneous MRS, M_{t+1}^c, is not necessarily zero and is denoted PE^j, where (dropping the time subscripts for brevity)

$$PE^j = E\left[M^c R^j\right] - 1. \tag{10}$$

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, since only the former are potentially nonzero if the model is true.

Under joint lognormality of C_{t+1}/C_t and returns, the pricing error may be written

$$PE^j = E\left[R^j\right] E\left[M^c\right] \exp\left\{\text{Cov}\left(m^c, r^j\right)\right\} - 1 \tag{11}$$

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

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

implying under joint lognormality,

$$\begin{aligned}
PE^j &= \frac{E\left[M^c\right]}{E\left[M^i\right]} \exp\left\{\text{Cov}\left(m^c, r^j\right) - \text{Cov}\left(m^i, r^j\right)\right\} - 1 \tag{12} \\
&= \frac{E\left[M^c\right]}{E\left[M^i\right]} \exp\left\{-\gamma_c \text{Cov}\left(\Delta c, r^j\right) + \gamma \text{Cov}\left(\Delta c^i, r^j\right)\right\} - 1. \tag{13}
\end{aligned}$$

How are the parameters and pricing errors distorted by using M_{t+1}^c to price assets in place of the true pricing kernel M_{t+1}^i? For $N > 2$ asset returns, it is not possible to give a intuitively appealing analytical expression for this distortion, although values can be obtained numerically. It is, however, possible to illustrate analytically the distortion in γ_c to a very close approximation, by focusing on log pricing errors and assuming that the risk-free rate is constant. In this case we can choose δ_c so that $E\left[M^i\right] = E\left[M^c\right]$, which insures that the

pricing error for the risk-free rate is zero.[14] While this is an approximation, it turns out to be very well satisfied in the data, since the Treasury-bill rate is extremely stable.[15] As a result, findings based on the full numerical solution are almost identical to those based on this approximation when the returns being priced are calibrated to match the means and volatilities of those in the data. We maintain this approximation purely for expositional purposes; the reader should be aware that exact results are very close.[16]

With this approximation in hand, the pricing error of the jth asset is now

$$PE^j = \exp\left\{-\gamma_c \text{Cov}\left(\Delta c, r^j\right) + \gamma \text{Cov}\left(\Delta c^i, r^j\right)\right\} - 1.$$

With δ_c set as just described and γ_c is chosen to minimize the sum of squared log pricing errors, $pe^j \equiv \log\left(1 + PE^j\right)$, as in GMM estimation using the identity weighting matrix, the resulting value of γ_c is given by

$$\widehat{\gamma}_c = \gamma \left(\frac{\sum_j \sigma_{cj}\sigma_{ij}}{\sum_j \sigma_{cj}^2}\right), \tag{14}$$

where $\sigma_{cj} \equiv \text{Cov}(\Delta c, r^j)$, $\sigma_{ij} \equiv \text{Cov}(\Delta c^i, r^j)$, and "$\sum_j$" indicates summation over all $j = s, f, 1, ..., N$ asset returns being priced. In the two-asset case, when only R^f and R^s are priced, this collapses to

$$\widehat{\gamma}_c = \gamma \left(\frac{\sigma_{is}}{\sigma_{cs}}\right). \tag{15}$$

Notice that, in the two asset case, this value of γ_c, along with the value of δ_c discussed above, insures that the log pricing errors for R_{t+1}^s and constant R^f are identically zero.[17] This follows

---

[14]Note that this does not imply that the risk-free rate puzzle is trivial, since δ_c is unrestricted and in particular can be chosen to be greater than unity if required to set the pricing error to zero.

[15]If M_t^i is the true pricing kernel, then $E\left[M_t^i\right] = E\left[1/R_t^f\right]$. Since we assume $E\left[M_t^i\right] = E\left[M_t^c\right]$, our assumption implies $E\left[M_t^c\right] = E\left[1/R_t^f\right]$, which prices the risk-free rate exactly if R_t^f is constant. It follows that the approximation error in pricing the risk-free rate is $E\left[1/R_t^f\right] - 1/E\left[R_t^f\right]$, which is -0.01 percent per annum. Therefore, the approximation implies we pick δ so that $E\left[M_t^i\right] = E\left[M_t^c\right]$, with the resulting pricing error in the risk-free rate of -0.01%.

[16]The calculations below are similar in spirit to those in Vissing-Jorgensen (1999), who shows how limited stock market participation biases estimates of relative risk aversion based on aggregate consumption. Vissing-Jorgensen's calculations presume heterogenous households rather than a representative-stockholder, as below.

[17]When there are only two asset returns, simple analytical expressions for the values of δ_c and γ_c that insure the pricing errors are identically zero can be obtained without assuming that the risk-free rate is constant. For a single risky asset return R_{t+1}^s and the risk-free return R_{t+1}^f, these values are given by

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

15

because, under lognormality, the log model is linear and the problem collapses to solving two linear equations in two unknowns. Therefore an erroneous pricing kernel based on aggregate consumption can always be found that unconditionally prices any 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 (9) to data will not correspond to any marginal investor'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.

Equation (14) shows that γ_c will be higher the higher is assetholder risk aversion, but that this relation depends on the statistical properties of aggregate consumption growth, individual consumption growth, and returns. Again, a more intuitively appealing expression can be obtained by focusing on the two-asset case. 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:

$$\Delta c_t = \beta \Delta c_t^i + \varepsilon_t^i, \tag{16}$$

where $\beta = \frac{\text{Cov}\left(\Delta c_t, \Delta c_t^i\right)}{\text{Var}\left(\Delta c_t^i\right)} \equiv \frac{\rho_{ci}\sigma_c}{\sigma_i}$. Here ρ_{ci} denotes the correlation between Δc_t and Δc_t^i. Using this decomposition, (15) can be re-written as

$$\widehat{\gamma}_c = \frac{\gamma}{\beta + \frac{\sigma_{\varepsilon^i s}}{\sigma_{is}}}, \tag{17}$$

where $\sigma_{\varepsilon^i s} \equiv \text{Cov}\left(\varepsilon_t^i, R_{t+1}^s\right)$. Now consider assets that are uncorrelated with ε_t^i, the component of aggregate consumption that is orthogonal to stockholder consumption. Any risky asset that is on the log mean-variance efficient frontier will be included in this category. In this case $\sigma_{\varepsilon^i s} = 0$ and therefore

$$\widehat{\gamma}_c = \frac{\gamma}{\beta} = \gamma \frac{\sigma_i}{\rho_{ci}\sigma_c}. \tag{18}$$

The formula tells us that 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 (9), but to do so, assetholder consumption must be more volatile than aggregate consumption and/or very weakly correlated correlated with it. Notice, however, that even

where $\sigma_{if} \equiv \text{Cov}\left(\Delta c^i, r^f\right)$, $\sigma_{cf} \equiv \text{Cov}\left(\Delta c, r^f\right)$, 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. Notice that, in equilibrium, $\widehat{\gamma}_c$ and $\widehat{\delta}_c$ 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.

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 (9) to data in the two-asset case. In that case, the only consequence of using aggregate per capita consumption in this setting is a bias in the estimated parameters $\widehat{\gamma}_c$ and $\widehat{\delta}_c$; there is no consequence for the Euler equation errors, which remain zero.

How do the pricing errors under lognormality compare with those estimated in the data when there are more asset returns? Figure 2 provides an answer for both the two- and eight-asset cases using actual historical return data. The "data" line plots RMSE/RMSR over a range of values for γ_c, that arise from choosing δ_c to minimize the sum of squared pricing errors in (10), which do not impose lognormality. The top panel plots for the two-asset case (these lines reproduce those in Figure 1), the bottom panel for the eight-asset case. The line labeled "lognormality" plots the RMSE/RMSR that arise from choosing δ_c to minimize the sum of squared pricing errors in (11), which impose lognormality. One way to interpret the "lognormal" line is to note that we can always find a pricing kernel $M_{t+1}^i = \exp\{\log(\delta) - \gamma \Delta c_{t+1}^i\}$, for some δ, γ, and normally distributed Δc_{t+1}^i, which along with (16) and a statistical model for log returns, such as

$$r_t^j = \alpha^j \Delta c_t^i + \eta_t^j,$$

generates a set of asset returns with the same means, variances and covariances with Δc_t as those in the historical data, and prices those asset exactly.[18] The dashed line labeled "lognormality" then gives the pricing errors that would arise from fitting M_{t+1}^c to data generated from this model.

Figure 2 shows that a lognormal model cannot explain the pricing errors in the data, especially when γ_c is large. When only two assets are priced (top panel), values for δ_c and γ_c can be found for which the pricing errors of the CRSP stock market return and the Treasury-bill rate are exactly zero, whereas this is not true in the data when no distributional assumptions are imposed. Similarly, the bottom panel shows that the lognormal model cannot match the magnitude of the Euler equation errors for the eight-asset case, increasingly so as γ_c rises.

We note that the results above hold for *any* pricing kernel M_{t+1}^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 (7). Referring to (12) 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, returns and m_t^i, whatever form the latter may take. As long as the true kernel M_{t+1}^i is jointly lognormally distributed with aggregate consumption and returns, then values for the

---

[18]This is done by choosing α^j to match the mean excess return for each asset, choosing $\text{var}(\eta^j)$ to match the volatility of each return, and choosing $\text{cov}(\eta, \varepsilon^i)$ to match the $\text{cov}(r^j, \Delta c)$ from the data.

discount factor relative risk aversion can be always be found such that the standard model generates identical (zero) unconditional asset pricing implications for two asset returns, and for which the root mean-squared pricing errors are much smaller than in the data (especially for large values of γ_c) for larger cross-sections of asset returns. These results suggest that models that are approximately lognormal will also have difficulty explaining the large pricing errors of the standard model.

3.2 Leading Asset Pricing Models

We now turn our attention to investigating how well leading asset pricing models explain the pricing errors of the standard model by examining the properties of model-generated data. All of the models generate predictions endogenously for a stock market return and a risk-free rate, and none imply that the pricing kernel is unconditionally jointly lognormally distributed with aggregate consumption growth and returns.[19] In addition, the Menzly, Santos, and Veronesi (2004) (MSV) model is multi-asset extension of the Campbell and Cochrane (1999) (CC) habit model; thus we can extend our analysis of nonlinear habit-based models to study multiple risky asset returns by applying the MSV framework. It is also straightforward to study the multi-asset properties of the long-run risk model of Bansal and Yaron (2004) (BY), since the cash-flows in BY are exogenously modeled. By contrast, the limited participation model of Guvenen (2003) (GUV) generates implications for only two asset returns, a single risky (stock market) return, and a risk-free return. Since cash-flows are endogenously determined by the properties of a general equilibrium setting in that model, the extension to multiple-assets is not trivial and would have to be developed. Thus, we focus only on the implications of the Guvenen model for R^s and R^f below.

3.2.1 Misspecified Preferences

We begin with the representative-agent models and consider three prominent representative agent models: the external habit-persistence models of Campbell and Cochrane (1999) and Menzly, Santos, and Veronesi (2004), and the long-run risk model of Bansal and Yaron (2004). 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. In what follows, we describe only the main features of each model, and refer the reader to the original article and the Appendix for details. Except where noted, our simulations use the baseline parameter values of each paper.

---

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

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

$$U = E \left\{ \sum_{t=0}^{\infty} \delta^t \frac{(C_t - X_t)^{1-\gamma} - 1}{1 - \gamma} \right\}, \qquad \gamma > 0 \tag{19}$$

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 heteroskedastic autoregressive processes. (The Appendix provides a 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 (see the Appendix). We denote as M_{t+1}^{CC} the specification of the stochastic discount factor corresponding to the Campbell-Cochrane model of X_t, and as M_{t+1}^{MSV} the specification of the stochastic discount factor corresponding to the MSV model of X_t. Both 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 (stock market claim) that pays aggregate consumption as its dividend. As above, the returns to these assets are denoted R_{t+1}^f, and R_{t+1}^s, respectively. 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.

The MSV model is a multi-asset extension of the CC model that generates implications for multiple risky securities, each distinguished by a distinct dividend process with dynamics characterized by fluctuations in the share it represents in aggregate consumption:

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

where n represents the total number of risky financial assets indexed by j, each paying a dividend D^j. Cross-sectional variation in unconditional mean returns across risky securities is governed by cross-sectional variation in the covariance between shares s_t^j and aggregate consumption growth Δc_t. In analogy to the empirical exercise, we create a model-generated cross-section of asset returns comprised of six risky securities plus the aggregate wealth

portfolio (stock market) return and the risk-free rate, for a total of 8 asset returns. The Appendix gives a detailed description of the stochastic process for the shares.

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

$$M_{t+1}^{BY} = \left(\delta \left(\frac{C_{t+1}}{C_t} \right)^{-\frac{1}{\psi}} \right)^{\alpha} R_{w,t+1}^{\alpha-1}, \tag{20}$$

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

$$\Delta c_{t+1} = \mu + x_t + \sigma_t \eta_{t+1} \tag{21}$$

$$\Delta d_{t+1} = \mu_d + \phi x_t + \rho_d \sigma_t u_{t+1}, \tag{22}$$

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

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

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 (22), rather than a claim to aggregate consumption, given by (21). We denote the return to this dividend claim R_{t+1}^s, since it corresponds the model's stock market return. BY calibrate the model so that x_t is very persistent, with a small unconditional variance. Thus, x_t captures long-run risk, 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 $\psi = 1.5$.

As for the MSV model, we can analyze the multi-asset implications of the BY model by considering risky securities, indexed by j, that are distinguished by their cash-flow processes:

$$\Delta d_{t+1}^j = \mu_d^j + \phi^j x_t + \rho_d^j \sigma_t u_{t+1}. \tag{23}$$

By considering a grid of values for ϕ^j, we create risky securities with different risk-premia, since this parameter governs the correlation of equilibrium returns with the stochastic discount factor. By altering ρ_d^j, we control the variance in the risky security returns, and μ_d^j controls the mean price-dividend ratio across risky assets.

For both the MSV and BY models, we choose parameters of the cash-flow processes to create a cross-section of asset returns that include a risk-free rate, an aggregate equity return, and six additional risky securities, or eight securities in total. For each model, we exactly replicate the original calibration of the risk-free return and aggregate equity return, themselves chosen to match the properties of these assets in U.S. data. For the six additional risky securities, we choose parameters of the cash-flow processes that allow us to come as close as possible to matching the spread spread in risk-premia found in the six size/book-market sorted portfolio returns in the data. For the BY model, we are able to generate a cross-section of returns that come very close to matching the historical spread. For example, the largest spread in average annualized returns is given by the difference between the portfolio in the smallest size and highest book-market category and the portfolio in the largest size and lowest book-market category, equal to about seven percent. Thus, we create six artificial returns for which the largest spread is 6.7 percent per annum. Constructing such returns for the MSV framework is more complicated, since the solutions for the multi-asset model hold only as an approximation (see the Appendix for the approximate relation). Unfortunately, we find that the approximation error in this model can be substantial under parameter values required to make the maximal spread as large as seven percent.[20] As a result, we restrict the parameter values to ranges that limit approximation error to reasonably small degrees. This still leaves us with a maximal spread of 4.5 percent per annum in the returns of the six artificial securities created.

To study the implications of these representative-agent models, we simulate a large time-series (e.g., 20,000 periods) from each model 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 fitted to data generated by these models. As in the historical data, the parameters γ_c and δ_c are chosen by GMM to minimize the Euler equation errors $E[M_{t+1}^c R_{t+1}^j] - 1$, $j = 1, ..., N$. We denote the parameters that minimize the GMM criterion as $\widehat{\delta}_c$ and $\widehat{\gamma}_c$. As in the historical data, we focus on the case of $N = 2$ asset returns (R_{t+1}^s and R_{t+1}^f), and the case of $N = 8$ asset returns, $(R_{t+1}^s, R_{t+1}^f, R_{t+1}^1, ..., R_{t+1}^6)$.

The results are presented in Table 3. First consider the CC and MSV habit models. For each model, we find the pricing errors that arise from fitting M_{t+1}^c to model-generated data are numerically zero, just as they are when the true habit pricing kernel is used. This

---

[20]Menzly, Santos, and Veronesi (2004) state that the approximation error is small for the parameters they employ, but it is not small for our parameters, which were chosen to mimic returns of the Fama-French portfolios.

result does not depend on the number of assets being priced; it is the same for the two-asset case and eight-asset case. 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, as measured by the root mean-squared pricing error. In the CC model, the values of δ_c and γ_c that set these pricing errors for R_{t+1}^s and R_{t+1}^f 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+1}^c to price assets arises. No distortion appears in the Euler equation errors themselves.

The conclusions for the Bansal-Yaron long-run risk model, also displayed in Table 3, are the same. Here we follow BY and simulate the model at monthly frequency, aggregate to annual frequency, and report the model's implications for pricing errors and parameter values. The monthly consumption data are time-aggregated to arrive at annual consumption, and monthly returns are continuously compounded to annual returns.[21] We find that δ_c is close to the true value, but γ_c is estimated to be about five times as high as true risk aversion. 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+1}^c to data generated by M_{t+1}^{BY}, the resulting Euler equation errors would be effectively zero, in contrast to what is found using historical data.[22] In the two-asset case the RMSE is zero to numerical accuracy, and it is 0.01% per annum in the eight-asset case.

3.2.2   Misspecified Consumption

We now consider the limited participation model of Guvenen (2003) (GUV). Like the representative agent models considered above, this model has 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. Here we suppose (1)

---

[21]The 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.

[22]For 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.

is fitted to data generated by this non-representative agent model in which asset prices are determined not by per capita aggregate consumption but rather by the consumption of stockholders.

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 denoted

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

where C_t^i is *stockholder* consumption, which by assumption is not the same as aggregate per capita consumption, δ is the subjective discount factor of the stockholder, and γ is the stockholder's relative risk aversion. 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 4, panel A shows that stockholder consumption growth is about two and a half times as volatile as aggregate consumption growth in the baseline model, and perfectly correlated with it. The model also implies that stockholder consumption is over four times as volatile as nonstockholder consumption growth, but the two are almost perfectly 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 marginal rate of substitution, M_{t+1}^{GUV}, is highly correlated with an aggregate consumption pricing kernel $M_{t+1}^c \equiv \delta_c (C_{t+1}/C_t)^{-\gamma_c}$, for a variety of values of δ_c and γ_c. Panel B of Table 4 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 $\widehat{\delta}_c = 0.99$ and $\widehat{\gamma}_c = 4.49$; thus, unlike the representative-agent models discussed above, this model does not explain the high estimated values of δ and γ obtained when fitting data to the standard consumption-based model. In both cases, the correlation between M_{t+1}^{GUV} and M_{t+1}^c is extremely high, 0.99.

Panel C of Table 4 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+1}^{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. To compute the pricing errors this model produces when M_{t+1}^c is employed to explain asset returns, we first set the parameters δ_c and γ_c equal to the stockholder's discount factor, δ, and risk aversion, γ, in the baseline parameterization. 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, but still small in magnitude compared to the data. By contrast, 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 risk-free return. By simply increasing γ by a factor of 2.5, to 4.5 from 2, the Guvenen model delivers a power utility pricing kernel using aggregate consumption that explains the historical mean return on the stock market and risk-free (Treasury bill) return just as well as the true pricing kernel based on stockholder consumption. Thus, the model does not explain the equity premium puzzle of Mehra and Prescott (1985), which is the puzzle that a high value of γ (greater than 10 according to Mehra and Prescott) is required to explain the equity premium when the power utility model is calibrated to aggregate consumption data. These results echo the empirical findings in Brav, Constantinides, and Geczy (2002), which suggest that limited stock market participation plays a minimal role in explaining the historical equity premium.

3.2.3 Misspecified Preferences and Misspecified Consumption

One possible reaction to the results above, is that we should take the representative agent nature of the CC, MSV and BY 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 limited-participation version of the MSV model and show that the conclusions are 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 individual assetholders, denoted

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

where X_{t+1}^i is the habit modeled as in MSV, not a function of C_t^i. We then compute two types of pricing errors. First, we assume that the data are generated by M_{t+1}^{MSVi} and compute the pricing errors that arise from using

$$M_{t+1}^{ch} \equiv \delta_c \left(\frac{C_{t+1} - X_{t+1}}{C_{t+1} - X_{t+1}} \right)^{-\gamma_c}$$

to price assets, where X_t is now computed from the MSV habit specification using aggregate consumption. Second, we assume the data are generated by M_{t+1}^{MSVi} and compute the pricing errors that arise from using

$$M_{t+1}^c \equiv \delta_c \left(C_{t+1}/C_t \right)^{-\gamma_c}$$

to price assets. This latter case is what we refer to as misspecified preferences and misspecified consumption; an econometrician who tried to fit M_{t+1}^c to asset return data would be employing both the wrong preferences and the wrong consumption measure. In both cases, δ_c and γ_c are chosen to minimize an equally-weighted sum of squared pricing errors of the assets under consideration, as with the historical data.

The results are presented in Table 5 for the exercise using M_{t+1}^{ch} and Table 6 for the exercise using M_{t+1}^c. We report the pricing errors for a range of parameter specifications. The standard deviation of asset-holder consumption growth is allowed to range from one times to five times as volatile as that of aggregate consumption growth, the correlation from -1.0 to 1.0. The pricing errors (as measured by RMSE/RMSR) are reported in the bottom subpanels. The top panel reports these errors for the two-asset case where only R_{t+1}^s and, R_{t+1}^f are priced; the bottom panel reports for the eight-asset case with six additional risky securities. For each parameter configuration, we also report the values $\widehat{\delta}_c$ and $\widehat{\gamma}_c$ that minimize the quadratic form $g_T(\gamma_c, \delta_c)$, as above.

Several aspects of Tables 5 and 6 are of interest. First, consider the two-asset case. Table 5 shows that the pricing errors that arise from using M_{t+1}^{ch} to price assets are always zero, even if assetholder consumption growth has very different properties from aggregate consumption growth. For example, Table 5 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. Notice, however, that the parameters δ_c and γ_c can deviate substantially from the true preference parameters of stockholders. This is similar to the lognormal example, in which the use of mismeasured consumption distorts preference parameters, but does not explain the large pricing errors generated by the standard consumption-based model.

Second, Table 6 shows that the same result holds if one uses M_{t+1}^c in place of the true pricing kernel M_{t+1}^i. Here the model used to explain asset returns is based *both* the wrong consumption measure and the wrong preferences. Nevertheless, values of δ_c and γ_c exist 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 5 where we have merely substituted the wrong consumption measure into the class of habit preferences, but the pricing errors are still zero.[23] These findings reinforce the conclusion that changing the pricing kernel does not necessarily change the pricing implications.

Third, results for the multi-asset case are qualitatively the same as those for two-asset case. For example, Table 6, bottom panel shows that the root mean-squared pricing error that arises from erroneously using M_{t+1}^c to price assets is a tiny fraction of the square-root of the average squared returns of the assets under consideration. The highest is 4% per annum. These numbers should be contrasted with the 33% figure obtained for a cross-section of 8 asset returns in U.S. data (Table 1). Moreover, the numbers in Table 6 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 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 (26). 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 6.

The results reveal a striking implication of leading asset pricing models: the unconditional

---

[23]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 the parameter α in the MSV habit specification (see the Appendix) at the same time as we adjust σ_i/σ_c so that the mean excess return $R^s - R^f$ 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.

pricing errors of the standard consumption-based model can be virtually identical to those using the true pricing kernel, even when (i) the true kernel has preferences different from the CRRA form of the standard model, (ii) the consumption of marginal assetholders behaves differently from per capita aggregate consumption, and (iii) the number of assets exceeds the number of free parameters to be estimated. This implies that the explanation for the high average pricing errors produced by the standard model has to be something more than limited participation and/or nonstandard preferences per se, since in many models parameter values can be found that allow the standard model to price cross-sections of assets almost as well as the true pricing kernel that generated the data.

3.2.4   Time Aggregated Consumption

What if the decision interval of households may be shorter than the data sampling interval, leading to time-aggregated consumption observations? 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 a variety of decision intervals. For example, we assume that agents make decisions quarterly but that the data sampling interval is annual. As above, we also allow for the possibility that aggregate consumption is a misspecified measure of assetholder consumption. For all models the essential results for the Euler equation errors remain small: values of δ_c and γ_c can always be found such that the unconditional pricing errors associated with using M_{t+1}^c to price assets are very small relative to the data, even when using time-averaged data. As an example, Table 7 shows results 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, although note that the results above for the BY model are already based on time-aggregate data.)[24] Most values of RMSE/RMSR are close to zero. The largest occurs for the eight asset case and is equal to 0.07, which occurs only if we assume stockholder consumption growth is negatively correlated with aggregate consumption growth. This is far smaller than the value of 0.33 found in the data. 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 are not crucial for explaining the large pricing errors of the standard model.

---

[24] Again, because of numerical error, these figures actually overstate the true relative pricing errors, since the RMSE for the true MSV stochastic discount factor is of the same order of magnitude as that for the CRRA model.

3.2.5 Finite Sample Pricing Errors

The results above are based on long samples of model-generated data, providing estimates of the population Euler equation errors. The estimates using historical data are based on a finite sample of 204 observations. In Table 8 we show that our main conclusions are robust to using samples equal in size to that of our historical dataset. The table reports the *maximum* RMSE/RMSR over 1,000 samples of size 204 that arises from fitting M_{t+1}^c to data generated from the relevant model. We do not report small-sample results for the eight-asset MSV model. The small sample behavior of the MSV model is problematic because the model is solved in continuous time and moreover holds only as an approximation for multiple risky securities. As a result, we find that small amounts of approximation error are compounded by discretization error in small samples and it is not possible to reduce these errors to reasonable levels unless the number of decisions within the period is almost infinite. Nevertheless, we are able to report the results for the two-asset case, since the solutions for the aggregate consumption claim and risk-free rate are not approximate. Table 8 shows that for the three representative agent models, CC, MSV, and BY, the maximum Euler equation errors are numerically zero, in both the two-asset and eight-asset cases. The Guvenen model produces a slightly higher maximum RMSE/RMSR in finite samples, equal to about 0.87% at an annual rate, but this is still well below the value of almost 50% found in historical data (Table 1). In short, the large empirical Euler equation errors of the standard model are not explained by small sample biases.

To summarize, the results above suggest that if the data on asset returns and consumption were generated by any of the leading models considered above, we would find zero Euler equation errors and the consequence of using the wrong pricing kernel would simply be incorrect estimates of δ and γ. We now move on to consider an alternative way to explain the large historical pricing errors in frictionless models, by further studying the potential roles of limited participation/incomplete markets. We saw above that when all variables are jointly lognormally distributed, the standard model does not in general generate the magnitude of pricing errors found in the data. Thus the next section considers models in which these variables are allowed to depart from joint 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 allowing for first-order Hermite expansions around the multivariate normal distribution. Since

many economic models are close to, if not exactly lognormal, this is advantageous because the leading term in the expansion is Gaussian, while higher-order terms accommodate deviations from normality. One caveat is that the distributions we consider cannot accommodate conditional heteroskedasticity or other forms of conditional temporal dependence. Allowing for such dependence along with arbitrary non-normalities would require the calibration of an infeasible number of Hermite parameters about which we have no information. 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 $y_t = (\Delta c_t, \Delta c_t^i, \Delta d_t)' \equiv (y_{1,t}, y_{2,t}, y_{3,t})'$, Δc_t is aggregate consumption growth, Δc_t^i is individual asset-holder consumption growth, and Δd_t is dividend growth of an aggregate stock market claim. We will consider 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 $h(y)$. A Hermite expansion is a polynomial in y times the standard Gaussian density. 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 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 first-order expansion but one that can nonetheless accommodate quite significant departures from normality. We investigate a large number of possible joint distributions by varying the parameters $a_0,...,a_3$. 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 estimate the parameters of $f(y)$ and $a(y)$.

For the equity claim, in equilibrium we must have

$$E_t \left[M_{t+1}^i \left(\frac{P_{t+1}}{D_{t+1}} + 1 \right) \frac{D_{t+1}}{D_t} \right] = \frac{P_t}{D_t}, \tag{24}$$

where $M_{t+1}^i \equiv \delta \left(C_{t+1}^i / C_t^i \right)^{-\gamma}$ is the true pricing kernel based on individual assetholder consumption, D_t denotes dividends paid at time t, and P_t is the end-of-period stock price at time t. For the risk-free rate, an analogous equation holds using the definition $R_{t+1}^f \equiv \left(E_t \left[M_{t+1}^i \right] \right)^{-1}$, but notice that since all variables are i.i.d., conditional expectations are just

the same as unconditional expectations and $h(y)$ can be used to compute (24) and the equilibrium risk-free rate. Also, the equilibrium price-dividend ratio is a constant, P/D, that satisfies

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

Given a distribution $h(y)$ and the equilibrium value for P/D, it is straightforward to compute the pricing errors associated with erroneously using $M_{t+1}^c \equiv \delta_c \left(C_{t+1}/C_t\right)^{-\gamma_c}$ to price assets. We assume the asset return data are generated by M_{t+1}^i and solve numerically for the values of δ_c and γ_c that minimize an equally-weighted sum of squared pricing errors that arise from using M_{t+1}^c to price assets.

For our numerical computations, 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 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 (2005). 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 and appears to depend on the horizon, sampling frequency, and sample size. It is estimated to be negative, equal to -0.000177 in the annual post-war data used by Lettau and Ludvigson (2005), 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 (2005)), and small positive values of the same order of magnitude, e.g., 0.000177. Finally, the parameters for asset-holder consumption and assetholder preferences are somewhat arbitrary since there is insufficient data available to measure these empirically. 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. The marginal distributions for Δc, Δc^i, and Δd for this example are presented in Figure 4. 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 Hermite parameters $a_o, ..., a_3$, are set to obtain the density shapes displayed in Figure 4.

For this particular joint distribution model, the RMSE that arises from erroneously using $M_{t+1}^c \equiv \delta_c \left(C_{t+1}/C_t \right)^{-\gamma_c}$ to price the two assets is 2.81%, close to the value 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. The latter is 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 corresponding standard deviation of Δd is a somewhat higher and its mean somewhat lower than the corresponding figures for the CRSP-VW return.

What features distinguish this example? First, notice from Figure 4 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. 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 density of aggregate consumption growth. Similarly, with probability 0.2, assetholder consumption can by 10% while dividend growth on the risky asset can grow 25%, again zero-probability events for aggregate consumption growth. It follows that simulations from such a distribution would deliver periods in which the joint behavior of M_{t+1}^c and returns would be quite different from the joint behavior of M_{t+1}^i and returns. Notice that assetholder consumption growth and returns are quite non-normal in this example, similar to findings in Brav, Constantinides, and Geczy (2002) that suggest higher-order moments of assetholder consumption growth have an important role in the pricing kernel.

Within the class of models we consider, how common is this example? To address this question, we evaluated pricing errors obtained from a wide grid (over 20,000 parameter combinations) 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 results. Two restrictions place limitations on the number of valid parameter combinations that can be considered. First, Ω must be positive semi-definite.[25] Second, the price-dividend ratio must be finite. Thus, risk-aversion cannot be too low if dividend growth is too high.

---

[25]Since the data suggest a weak correlation between aggregate consumption growth and dividend growth, this requires that the correlation between Δc_t and Δc_t^i be relatively small.

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 9 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+1}^c to price assets in place of M_{t+1}^i are, to numerical accuracy, zero. For example, the kurtosis of the marginal distribution of Δc_t is often greater than 11, and the skewness greater than 4. The values of γ_c and δ_c that drive the pricing errors to zero vary, 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 asset-holder consumption growth is much more volatile than aggregate consumption growth or when it is not highly correlated with it, as suggested by (18). 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 (15), where the expression in (15) collapses to $\gamma_c = \gamma \sigma_{id}/\sigma_{cd}$ in this model, so that γ_c is negative when σ_{cd} is negative.

Figure 5 provides a graphical description of two of the perturbed densities that created the output in Table 9. Notice that the shapes can differ considerably from Gaussian and yet values for γ_c and δ_c can still be found for which M_{t+1}^c prices assets just as well as the true kernel based on assetholder consumption. The densities in the left-hand column are bimodal for Δc^i and Δd, while aggregate consumption is close to normal. This is similar to the example above (Figure 4), which does deliver large pricing errors, but unlike that case the negative mass points are much smaller relative to the positive mass points. Also, in Figure 4 assetholder consumption growth has a higher mean than aggregate consumption growth, whereas in Figure 5 it has about the same mean. By contrast, the densities in the right-hand column of Figure 5 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.

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

$$\Delta d^j = \lambda^j \Delta c_t^i + \varepsilon_t^j, \qquad j = 1, ... N,$$

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 $y_t = \left(\Delta c_t, \Delta c_t^i, \Delta d_t^1, ..., \Delta d_t^N\right)'$ is assumed to be i.i.d. The "leverage" parameter λ^j controls the covariance of each asset return with the 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. 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.

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 $\text{Var}(\varepsilon_t^j)$. Equilibrium returns R_{t+1}^j are computed as described in the previous subsection for the two-asset case. As above, we assume the data are generated by $M_{t+1}^i \equiv \delta \left(C_{t+1}^i / C_t^i \right)^{-\gamma}$ and search numerically for values for δ_c and γ_c that minimize the Euler equation errors associated with using $M_{t+1}^c \equiv \delta_c \left(C_{t+1}/C_t \right)^{-\gamma_c}$ to price assets. The pricing errors are summarized by reporting the root mean-squared pricing error as a fraction of the root mean squared returns of the assets under consideration (RMSE/RMSR).

Table 10 presents both the maximum and average values of RMSE/RMSR obtained over a large grid search of distributional parameter values, including the special case of joint lognormality. The average pricing errors are often very small, indeed close to zero, even for significant perturbations from joint lognormality. We find a small number of cases in which the RMSE/RMSR is as large as 10 percent. Nevertheless, the 10 percent magnitude is still significantly smaller than in the data, and these cases are relatively rare, occurring in less than 0.2% of the parameter permutations. Most non-normal models we considered imply that the wrong pricing kernel based on aggregate consumption delivers tiny pricing errors even when the joint distribution of Δc_t, Δc_t^i, and returns are significantly non-normal. This suggests that the explanation for the large pricing errors of the standard representative agent model must be more than limited participation *per se*. The joint distribution of assetholder, aggregate consumption and returns has to be of a particular form, and it is that form that must be the central part of the story.

4   Conclusion

We view the evidence presented above as a convenient diagnostic for what remains missing in modern-day asset pricing theories designed to remedy shortcomings of the standard representative agent, consumption-based asset pricing model. In this paper we emphasize one shortcoming of the standard model that provides a margin upon which it fails overwhelmingly: its inability to explain the average returns on cross-sections of risky assets. 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 constitute a puzzle that is at least as damning as other, more well known, puzzles commonly emphasized when studying calibrated models.

Are prominent modifications to the standard model capable of explaining its mispricing? If so, then an econometrician who fit the standard model to data generated from leading asset pricing models should find large unconditional asset pricing errors, as in the historical data. Alas, we find that new pricing kernels do not necessarily generate new pricing implications.

Instead, we find that parameter values can often be found that imply the standard model has virtually 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 one risky and one risk-free asset, and for explaining larger cross-sections of risky returns. Moreover, some leading models imply that the standard consumption-based model is equally capable of explaining asset returns even when it is based both on the wrong consumption measure (aggregate consumption instead of individual assetholder consumption) and on the wrong model of underlying preferences (CRRA instead of habit or recursive preferences). The asset pricing models we explore *can*, in many cases, explain why an econometrician obtains implausibly high estimates of δ and γ when fitting the standard consumption-based model to historical data. 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 should account for these empirical regularities.

Inability to account for these empirical regularities cannot be uncovered by studying calibrated models or by procedures that rely solely on a model's first-order conditions for estimation and testing. 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 assess whether its key properties are matched by the simulated data of theoretical models.

Intuitively, how is it that asset pricing models capable of explaining a host consumption-based asset pricing puzzles are incapable of explaining the large unconditional Euler equation errors of the standard model? In thinking about this, it is helpful to consider the equity premium puzzle as an 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)).[26] 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 these lines? We considered examples of limited participation/incomplete markets models in which non-normalities are important, a finding also hinted at by the work of Brav, Constantinides, and Geczy (2002). But we also found that it is insufficient for assetholder consumption

---

[26] For the sake of argument, ignore the risk-free rate puzzle and other problems this resolution would leave.

to merely behave differently from aggregate consumption. This suggests that careful attention to the joint properties of the pricing kernel, aggregate consumption, and returns is crucial for explaining the mispricing of the standard paradigm in frictionless models. Alternatively, 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), may present 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 candidates include any modifications to the standard model that would make unconditional Euler equations more difficult to satisfy, 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

$$U = E\left\{\sum_{t=0}^{\infty} \delta^t \frac{(C_t^i - X_t^i)^{1-\gamma} - 1}{1 - \gamma}\right\}, \qquad \gamma > 0 \tag{25}$$

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:

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

where $\overline{Y}$ is the long-run mean of Y and k, α, and λ are parameters, calibrated as in MSV. Here $\Delta c_t \equiv \log(C_{t+1}/C_t)$, 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:

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

and compute equilibrium asset returns based on the stochastic discount factor $M_{t+1}^{MSVi} = \delta \left(C_{t+1}^i/C_t^i \right)^{-\gamma} \left(Y_t^i/Y_{t+1}^i \right)^{-\gamma}$. 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 $M_{t+1}^{ch} \equiv \delta_c \left(C_{t+1}/C_t \right)^{-\gamma_c} \left(Y_t^c/Y_{t+1}^c \right)^{-\gamma_c}$ in place of M_{t+1}^{MSVi} to price assets, where δ_c and γ_c are chosen freely to fit the data, and where Y_t^c follows the process

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

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,

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

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

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

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 $j = 1, ..., n$. 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 $\boldsymbol{v}^j$ as well as the individual share processes

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

We choose the parameters ϕ^j, $\bar{s}^j$, and $\boldsymbol{v}^j$, 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

$$\frac{P_t^j}{D_t^j} \approx a_0^j + a_1^j S_t + a_2^j \frac{\bar{s}^j}{s_t^j} + a_3^j \frac{\bar{s}^j}{s_t^i} S_t, \tag{26}$$

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

$$R_{t+1}^i = \left(\frac{P_{t+1}^j/D_{t+1}^j + 1}{P_t^j/D_t^j} \right) \exp \left(\Delta d_{t+1}^j \right). \tag{27}$$

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

$$U = E \left\{ \sum_{t=0}^{\infty} \delta^t \left\{ (1-\delta) C_t^{\frac{1-\gamma}{\alpha}} + \delta \left(E_t U_{t+1}^{1-\gamma} \right)^{\frac{1}{\alpha}} \right\}^{\frac{\alpha}{1-\gamma}} \right\}, \tag{28}$$

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

$$M_{t+1}^{BY} = \left(\delta \left(\frac{C_{t+1}}{C_t} \right)^{-\frac{1}{\psi}} \right)^{\alpha} R_{w,t+1}^{\alpha-1}, \tag{29}$$

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.

Brav, A., G. M. Constantinides, and C. C. Geczy (2002): "Asset Pricing with Heterogeneous Consumers and Limited Participation: Empirical Evidence," *Journal of Political Economy*, 110(4), 793–824.

Breeden, D., M. Gibbons, and R. Litzenberger (1989): "Empirical Tests of the Consumption-Oriented CAPM," *Journal of Finance*, 44, 231–262.

Campbell, J. Y. (1996): "Understanding Risk and Return," *Journal of Political Economy*, 104(2), 298–345.

——— (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.

——— (2000): "Explaining the Poor Performance of Consumption-Based Asset Pricing Models," *Journal of Finance*, 55(6), 2863–2878.

Cochrane, J. H. (1996): "A Cross-Sectional Test of an Investment-Base Asset Pricing Model," *Journal of Political Economy*, 104, 572–621.

——— (2005): *Asset Pricing*. Princeton University Press, Princeton, NJ.

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.

——— (1993): "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics*, 33, 3–56.

FERSON, W. E., AND G. M. CONSTANTINIDES (1991): "Habit Persistence and Durability in Aggregate Consumption," *Journal of Financial Economics*, 29, 199–240.

FERSON, W. E., AND J. J. MERRICK (1987): "Non-stationary and Stage of the Business Cycle Effects in Consumption-Based Asset Pricing Relations," *Journal of Financial Economics*, 18, 127–146.

GALLANT, A. R., AND G. TAUCHEN (1989): "Seminonparametric Estimation of Conditionally Constrained Heterogeneous Processes: Asset Pricing Applications," *Econometrica*, 57(5), 1091–1120.

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.

——— (1997): "Assessing Specific Errors in Stochastic Discount Factor Models," *Journal of Finance*, 52, 557–590.

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, Kellog 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.

——— (1996): "The Equity Premium: It's Still A Puzzle," *Journal of Economic Literature*, 34, 42–68.

LETTAU, M., AND S. C. LUDVIGSON (2001): "Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia are Time-Varying," *Journal of Political Economy*, 109(6), 1238–1287.

——— (2005): "Expected Returns and Expected Dividend Growth," *Journal of Financial Economics*, 76, 583–626.

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.

MANKIW, N. G., AND M. SHAPIRO (1986): "Risk and Return: Consumption Beta versus Market Beta," *Review of Economics and Statistics*, 68, 452–59.

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.

MENZLY, L., T. SANTOS, AND P. VERONESI (2004): "Understanding Predictability," *Journal of Political Economy*, 112(1), 1–47.

VISSING-JORGENSEN, A. (1999): "Limited Stock Market Participation and the Equity Premium Puzzle," Unpublished manuscript, Northwestern University, Kellog School of Management.

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

Table 1: Pricing Errors with CRRA Preferences

Assets	$\widehat{\delta}$	$\widehat{\gamma}$	RMSE (in %)	RMSE/RMSR	$p\,(W=I)$	$p\,(W=S^{-1})$
R^s, R^f	1.41	89.78	2.71	0.48	N/A	N/A
R^s, R^f, 6 FF	1.39	87.18	3.05	0.33	0.00	0.00
Excluding Periods with low Consumption Growth						
R^s, R^f	2.55	326.11	0.73	0.13	N/A	N/A
R^s, R^f, 6 FF	2.58	356.07	1.94	0.21	0.00	0.00

Notes: This table reports the minimized annualized postwar data pricing error for CRRA preferences. The preference parameters $\widehat{\delta}_c$ and $\widehat{\gamma}_c$ are chosen to minimize the mean square pricing error for different sets of returns: $min_{\delta_c,\gamma_c}\left[g(\delta_c,\gamma_c)'Wg(\delta_c,\gamma_c)\right]$ where $g(\delta_c,\gamma_c) = E[\delta_c(C_t/C_{t-1})^{-\gamma_c}\boldsymbol{R}_t - 1]$. 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. The table also reports results when the periods with the lowest six consumption growth rates are eliminated. The table reports estimated $\widehat{\delta},\widehat{\gamma}$ and the minimized value of RMSR/RMSRR where RMSE is the square root of the average squared pricing error and RMSR is the square root of the averaged squared returns of the assets under consideration for $W = I$. The last two colums report χ^2 p-values for tests for the null hypothesis that pricing errors are jointly zero for $W = I$ and $W = S^{-1}$ where S is the spectral density matrix at frequency zero. The data span the period 1951Q4 to 2002Q4.

Table 2: Low Consumption Growth Periods

Quarter	NBER Recession Dates	$C_t/C_{t-1} - 1$	R_t^s	R_t^f
1980Q02	80Q1-80Q3	-1.28	16.08	3.59
1990Q04	90Q3-91Q1	-0.87	8.75	2.16
1974Q01	73Q4-75Q1	-0.85	-1.26	2.37
1958Q01	57Q3-58Q2	-0.84	7.03	0.65
1960Q03	60Q2-61Q1	-0.64	-4.93	0.67
1953Q04	53Q1-54Q2	-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: Pricing Errors

Model	$\widehat{\delta}_c$	$\widehat{\gamma}_c$	RMSE/RMSR (R^s, R^f)	RMSE/RMSR (8 assets)
CC Habit	1.28	57.48	0.00	N/A
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 $\widehat{\delta}_c$ and $\widehat{\gamma}_c$ are chosen to minimize the mean square pricing error $min_{\delta_c,\gamma_c} \left[g(\delta_c, \gamma_c)' g(\delta_c, \gamma_c) \right]$ where $g(\delta_c, \gamma_c) = E[\delta_c(C_t/C_{t-1})^{-\gamma_c} \boldsymbol{R}_t - 1]$. RMSR is the square root of the averaged squared returns of the assets under consideration. RMSE is the square root of the average squared pricing error. Pricing errors are computed from simulations with 10,000 observations.

Table 4: 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 Factors

	$M_t^i(0.99, 2.00)$	$M_t^c(0.99, 2.00)$	$M_t^c(0.99, 4.49)$
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 Guvenen'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 5: Limited Participation Habit Model Estimated with Aggregate Consumption 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
	2 Assets: R^s, R^f					
	$\widehat{\delta}_c$					
1	0.84	0.66	0.25	0.39	0.84	0.96
2	0.85	0.65	0.22	0.46	0.83	0.96
5	0.84	0.66	0.23	0.46	0.82	0.96
	$\widehat{\gamma}_c$					
1	-1.69	-2.91	-4.89	5.11	2.35	0.84
2	-1.59	-3.09	-5.46	5.08	2.34	1.14
5	-1.75	-2.99	-5.96	4.88	2.26	1.11
	RMSE/RMSR					
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
	8 Assets					
	$\widehat{\delta}_c$					
1	0.84	0.65	0.27	0.42	0.84	0.96
2	0.84	0.65	0.27	0.40	0.83	0.96
5	0.85	0.63	0.23	0.36	0.81	0.96
	$\widehat{\gamma}_c$					
1	-1.71	-2.93	-5.21	5.08	2.28	1.02
2	-1.73	-2.96	-5.17	5.24	2.37	1.03
5	-1.76	-3.07	-5.46	5.39	2.51	1.18
	RMSE/RMSR					
1	0.04	0.03	0.03	0.03	0.03	0.02
2	0.05	0.03	0.04	0.03	0.02	0.02
5	0.05	0.04	0.04	0.03	0.02	0.02

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 $\widehat{\delta}_c$ and $\widehat{\gamma}_c$ are chosen to minimize the mean square pricing error $min_{\delta_c, \gamma_c} \left[g(\delta_c, \gamma_c)' W g(\delta_c, \gamma_c) \right]$ where $g(\delta_c, \gamma_c) = E[M_t^{ch} \boldsymbol{R}_t - 1]$, $M_t^{ch} = \delta_c \left(\frac{C_t}{C_{t-1}} \frac{Y_{t-1}}{Y_t} \right)^{-\gamma_c}$. 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 $W = I$. In the top panel, $\boldsymbol{R} = [R^s, R^f]'$, in the bottom panel $\boldsymbol{R}$ includes the return of the market R^s, the riskfree rate R^f and the returns of six individual assets. RMSR is the square root of the averaged squared returns of the assets under consideration. RMSE is the square root of the average squared pricing error. The weighting matrix W is the identity matrix.

Table 6: Limited Participation Habit Model Estimated with Aggregate Consumption 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
	2 Assets: R^s, R^f					
	$\widehat{\delta}_c$					
1	0.51	0.24	0.03	5.27	2.69	1.61
2	0.52	0.24	0.03	5.20	2.75	1.83
5	0.48	0.23	0.03	4.94	2.81	1.79
	$\widehat{\gamma}_c$					
1	-30.71	-60.15	-128.80	127.03	58.59	27.93
2	-29.22	-61.24	-132.02	117.99	61.69	33.28
5	-33.48	-64.30	-131.01	117.94	64.43	32.56
	RMSE/RMSR					
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
	8 Assets					
	$\widehat{\delta}_c$					
1	0.50	0.24	0.04	5.44	2.76	1.74
2	0.50	0.23	0.04	5.60	2.80	1.74
5	0.48	0.21	0.03	5.74	2.94	1.85
	$\widehat{\gamma}_c$					
1	-30.83	-61.99	-123.23	124.24	61.51	31.21
2	-31.69	-62.76	-124.21	126.92	62.34	31.22
5	-33.73	-67.43	-134.53	133.41	65.50	34.11
	RMSE/RMSR					
1	0.03	0.03	0.03	0.03	0.04	0.03
2	0.04	0.03	0.03	0.03	0.03	0.03
5	0.03	0.03	0.04	0.03	0.03	0.04

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 $\widehat{\delta}_c$ and $\widehat{\gamma}_c$ are chosen to minimize the mean square pricing error $min_{\delta_c,\gamma_c}\left[g(\delta_c,\gamma_c)'Wg(\delta_c,\gamma_c)\right]$ where $g(\delta_c,\gamma_c) = E[M_t^c \boldsymbol{R}_t - 1], M_t^c = \delta_c(\frac{C_t}{C_{t-1}})^{-\gamma_c}$. C_t is aggregate consumption, R^s is the return of equity, R^f is the riskfree rate, and $W = I$. $\boldsymbol{R}$ includes the return of the market R^s, the riskfree rate R^f and the returns of six individual assets. RMSR is the square root of the averaged squared returns of the assets under consideration. RMSE is the square root of the average squared pricing error. The weighting matrix W is the identity matrix.

Table 7: Limited Participation Habit Model Estimated with Aggregate Consumption CRRA SDF: 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
			2 Assets: R^s, R^f			
			$\widehat{\delta}_c$			
1	0.53	0.22	0.07	4.30	2.48	1.56
2	0.55	0.29	0.06	6.05	2.64	1.77
5	0.49	0.27	0.06	5.45	2.75	1.77
			$\widehat{\gamma}_c$			
1	-28.74	-64.81	-104.82	104.40	53.96	26.56
2	-26.32	-53.79	-110.70	129.91	59.36	31.78
5	-32.16	-58.47	-108.09	129.49	63.45	32.38
			RMSE/RMSR			
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
			8 Assets			
			$\widehat{\delta}_c$			
1	0.49	0.24	0.04	5.51	2.78	1.75
2	0.49	0.23	0.04	5.57	2.86	1.76
5	0.48	0.21	0.03	5.63	2.96	1.83
			$\widehat{\gamma}_c$			
1	-32.13	-61.76	-122.15	126.00	61.95	31.39
2	-32.02	-62.97	-125.44	127.60	63.43	31.83
5	-33.53	-66.57	-135.09	131.36	65.84	33.66
			RMSE/RMSR			
1	0.03	0.03	0.06	0.03	0.04	0.04
2	0.04	0.03	0.06	0.04	0.03	0.04
5	0.04	0.03	0.07	0.03	0.04	0.04

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 $\widehat{\delta}_c$ and $\widehat{\gamma}_c$ are chosen to minimize the mean square pricing error $min_{\delta_c,\gamma_c}\left[g(\delta_c,\gamma_c)'Wg(\delta_c,\gamma_c)\right]$ where $g(\delta_c,\gamma_c) = E[M_t^c \mathbf{R}_t - 1], M_t^c = \delta_c(\frac{C_t}{C_{t-1}})^{-\gamma_c}$. $\mathbf{R}$ includes the return of the market R^s, the riskfree rate R^f and the returns of six individual assets. RMSR is the square root of the averaged squared returns of the assets under consideration. RMSE is the square root of the average squared pricing error. The weighting matrix W is the identity matrix. C_t is aggregate consumption, R^s is the return of equity, R^f is the riskfree rate, and $W = I$. The model is simulated on a weekly frequency. The pricing errors are computed using the growth rate of annual consumption and compounded annual returns.

Table 8: Small Sample Pricing Errors

Model	max. RMSE/RMSR (2 Assets)	max. RMSE/RMSR (8 Assets)
CC Habit	0.01	N/A
MSV Habit	0.00	N/A
BY LR Risk	0.00	0.01
Guvenen Lim. Part.	0.87	N/A

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), Bansal and Yaron's long run risk model (BY LR Risk) and Guvenen's limited participation model. The preference parameters $\widehat{\delta}_c$ and $\widehat{\gamma}_c$ are chosen to minimize the mean square pricing error $min_{\delta_c,\gamma_c} \left[g(\delta_c, \gamma_c)' W g(\delta_c, \gamma_c) \right]$ where $g(\delta_c, \gamma_c) = E[\delta_c(C_t/C_{t-1})^{-\gamma_c} \boldsymbol{R}_t - 1]$. $\boldsymbol{R} = [R^s, R^f]'$ and $W = I$. "max. RMSE/RMSR" report the maximum absolute value of RMSE/RMSR from 5,000 simulations with 204 observations.

Table 9: Lim. Partic./Inc. Markets Pricing Errors for Stock Return and Risk-Free Rate: Hermite Densities

Cov(Δc,Δd)=0.00017

γ	δ	ρ(Δc,Δc')	ρ(Δc',Δd)	σ(i)/σ(c)	µ(Δc')/µ(Δc)	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	36.211	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(Δc,Δd)=-0.00017

γ	δ	ρ(Δc,Δc')	ρ(Δc',Δd)	σ(i)/σ(c)	µ(Δc')/µ(Δc)	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	-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(Δc,Δd)=0.00017

γ	δ	ρ(Δc,Δc')	ρ(Δc',Δd)	σ(i)/σ(c)	µ(Δc')/µ(Δc)	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	35.488	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(Δc,Δd)=-0.00017									
γ	δ	ρ(Δc,Δc')	ρ(Δc',Δd)	σ(i)/σ(c)	μ(Δc')/μ(Δc)	γc	δc	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(Δc,Δd)=0.00017									
γ	δ	ρ(Δc,Δc')	ρ(Δc',Δd)	σ(i)/σ(c)	μ(Δc')/μ(Δc)	γc	δc	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(Δc,Δd)=-0.00017									
γ	δ	ρ(Δc,Δc')	ρ(Δc',Δd)	σ(i)/σ(c)	μ(Δc')/μ(Δc)	γc	δc	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 persumed 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 10: Lim. Partic./Inc. Markets Pricing Errors in a Larger Cross-Section: Hermite Densities

Distribution	γ	δ	$\rho(\Delta c, \Delta c^i)$	$\sigma(\Delta c^i)/\sigma(\Delta c)$	$\mu(\Delta c^i)/\mu(\Delta c)$	$Cov(\Delta c, \Delta d)$	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 labled "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 labled "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 labled "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 persumed 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.

Figure 1: Pricing Errors for CRRA Preferences: Excess Returns
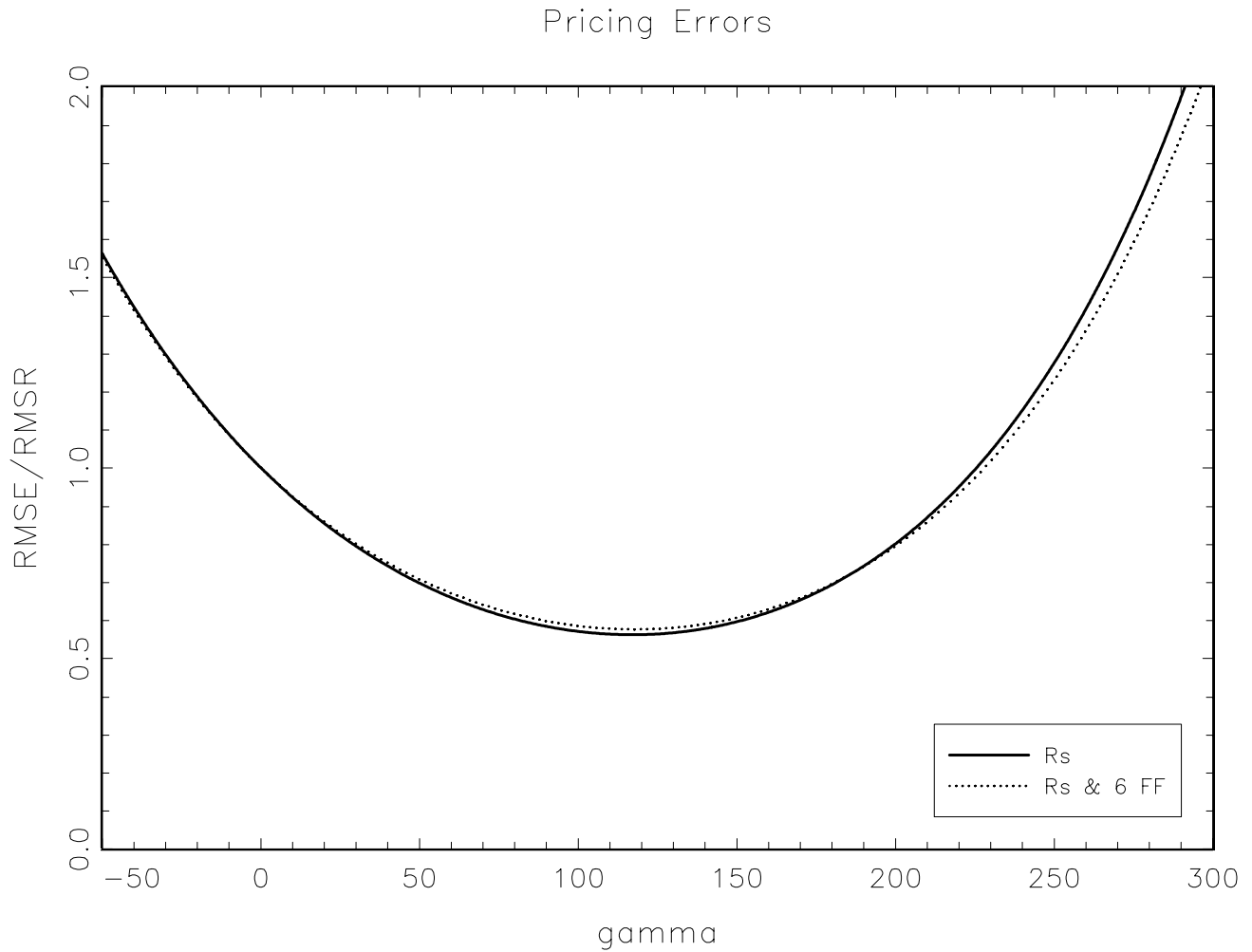
Pricing Errors

Notes: The figure plots RMSE/RMSR as a function of γ for excess returns. The pricing errors are PE $= E\left[(C_{t+1}/C_t)^{-\gamma}(\boldsymbol{R}_{t+1} - R^f_{t+1})\right]$. The solid line shows RMSE/RMSR for $\boldsymbol{R} = R^s$, the dotted line shows RMSE/RMSR for $\boldsymbol{R} = (R^s, 6\text{ FF})$.
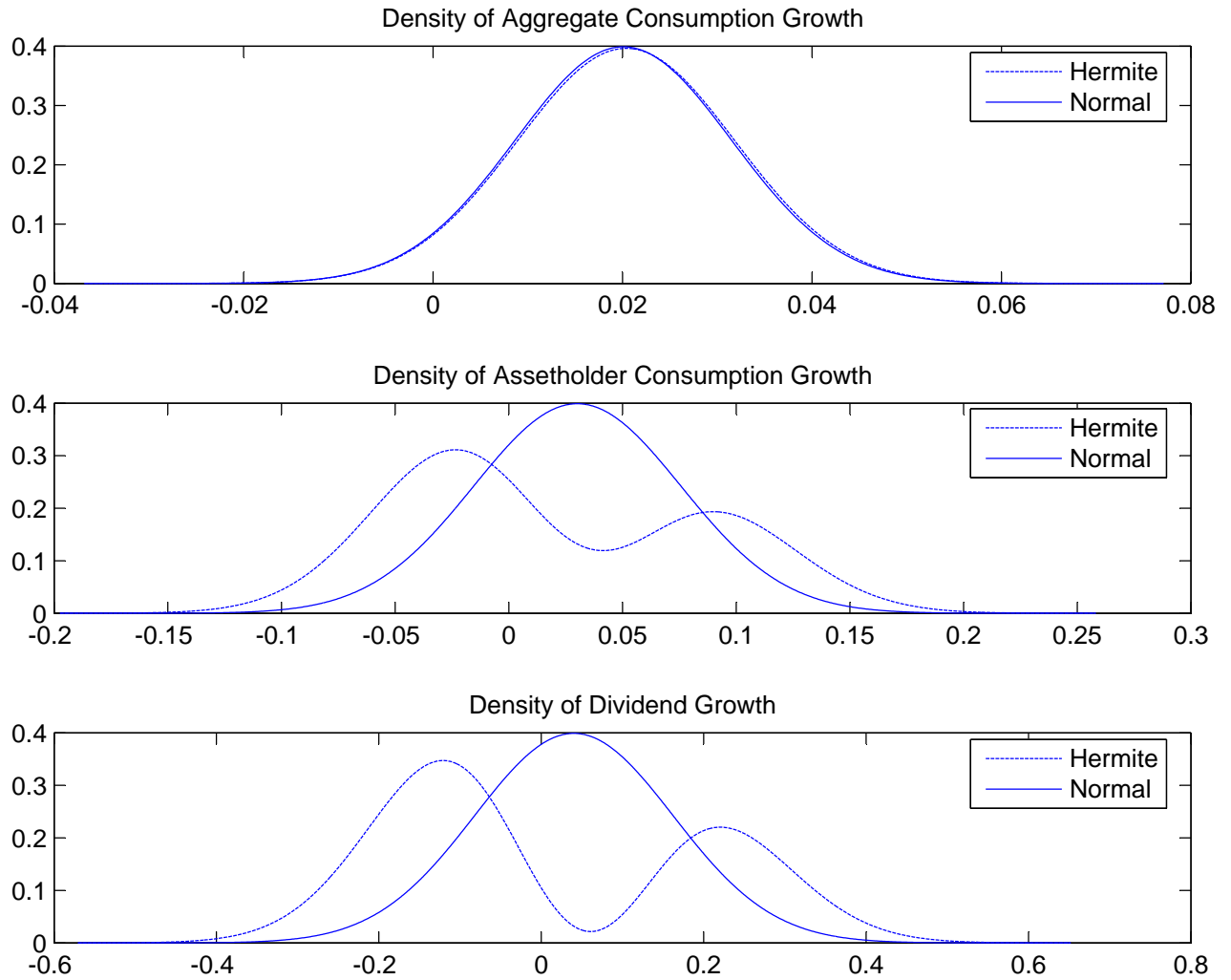
Figure 2: Pricing Errors: With and Without Lognormality

RMSE/RMSR: 2 Assets

RMSE/RMSR: 8 Assets
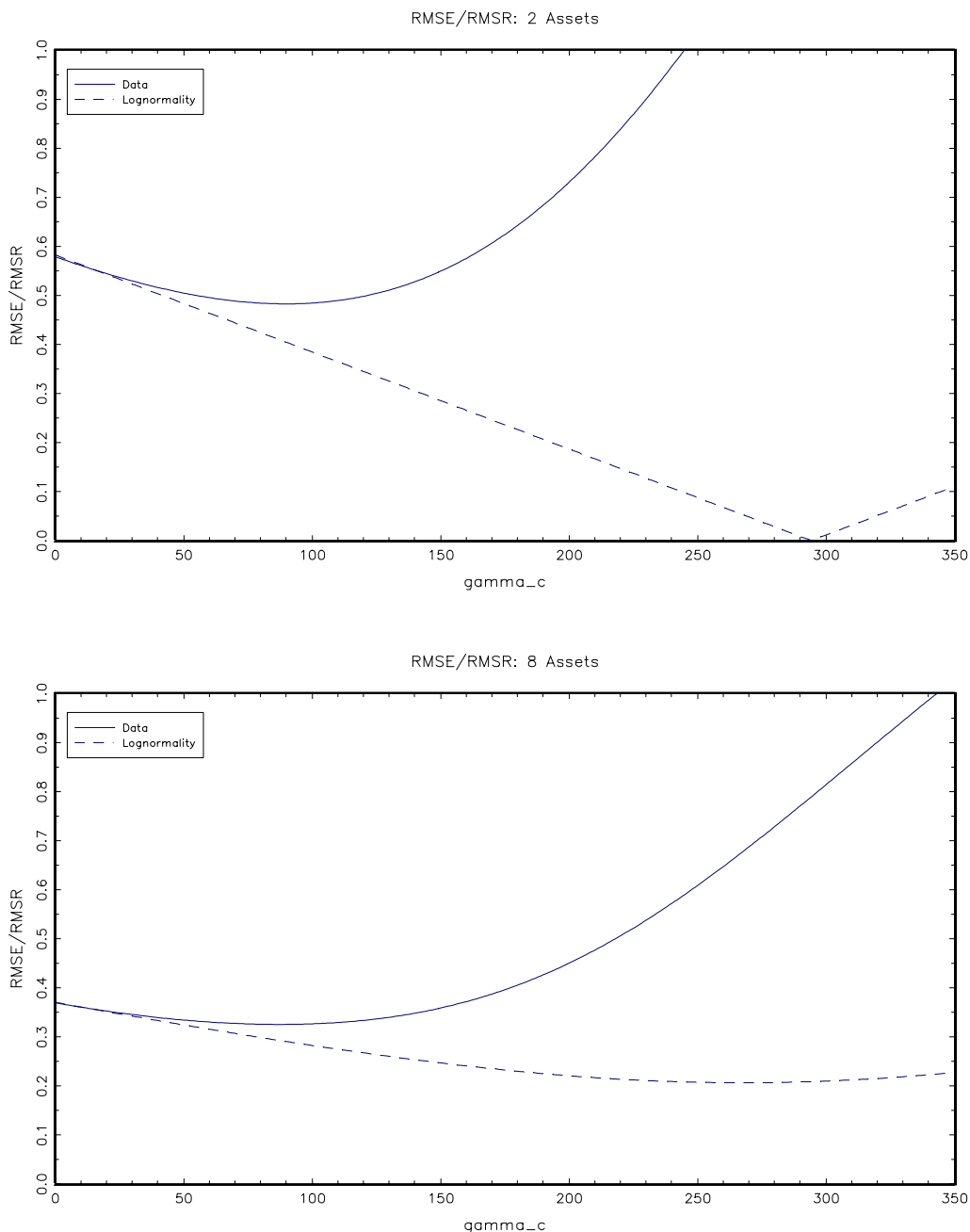
Notes: This figure plots RMSE/RMSR with and without the assumption of joint lognormality as a function of γ_c. δ_c is chosen to minimize the RMSE for each value of γ_c. The top panel shows the case for $\boldsymbol{R} = (R^s, R^f)$, in the bottom panel $\boldsymbol{R} = (R^s, R^f, 6 \text{ FF})$. The pricing error for asset j without assuming lognormality is $\text{PE}^j = \delta_c E \left[\exp \left\{ -\gamma_c \Delta c + r^j \right\} \right] - 1$. Under the assumption of joint lognormality, the pricing error is $\text{PE}^j = \delta_c \exp \left\{ -\gamma_c E \Delta c + \gamma_c^2 \sigma_c^2/2 + E r^j + \sigma_r^2/2 - \gamma_c \text{Cov}(\Delta c, r^j) \right\} - 1$.

Figure 3: QQ Plots
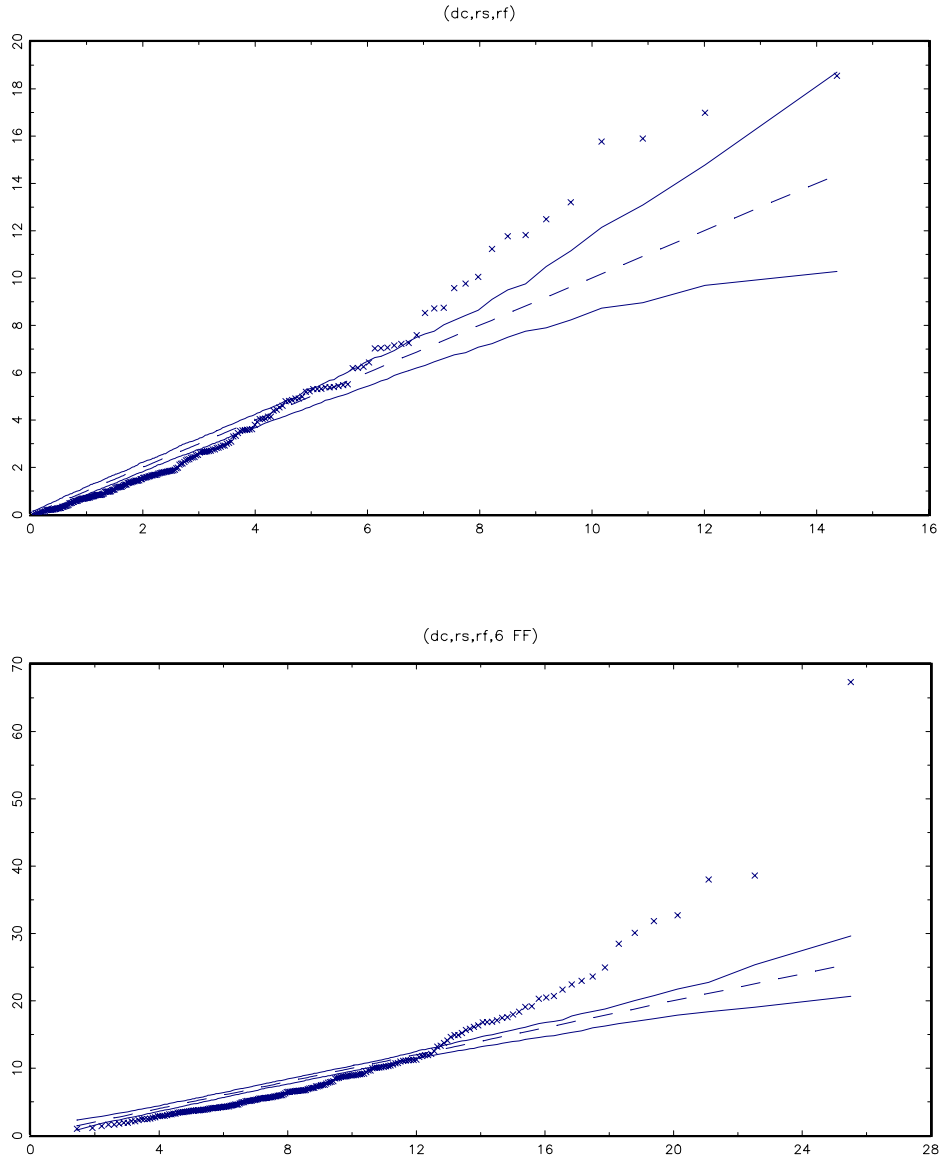
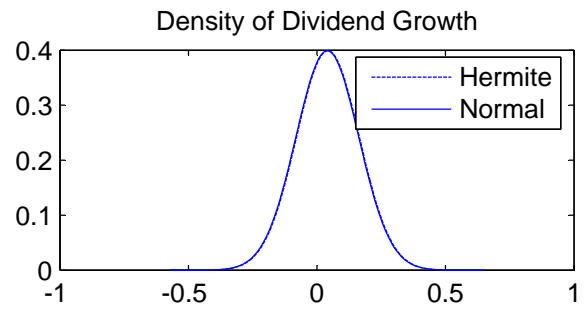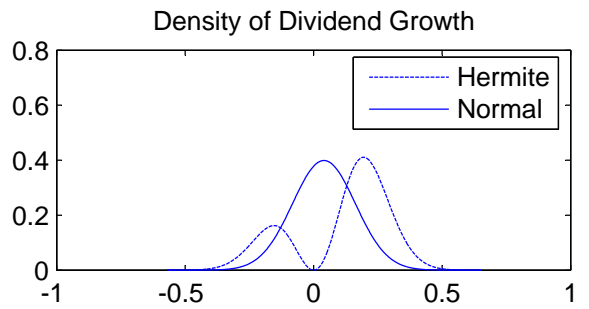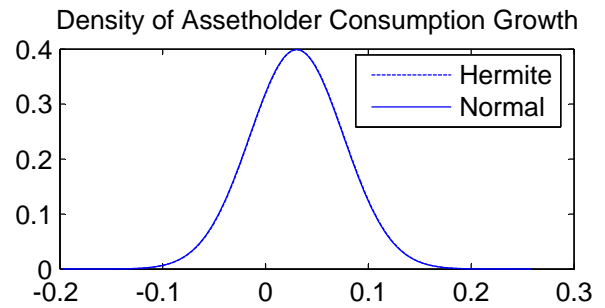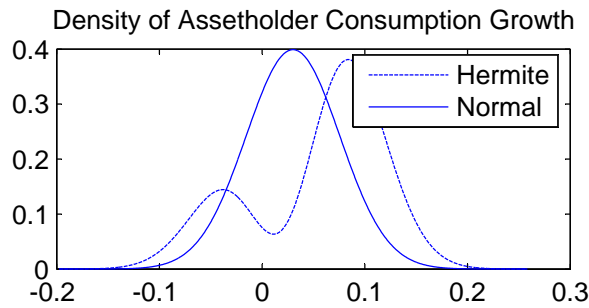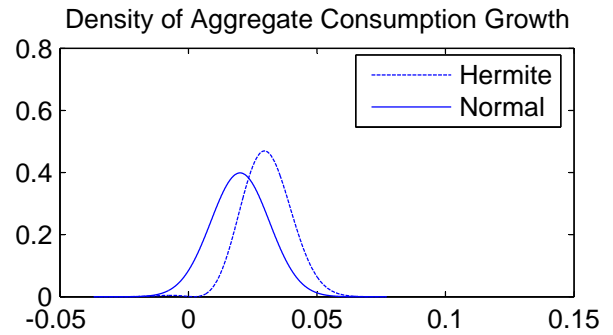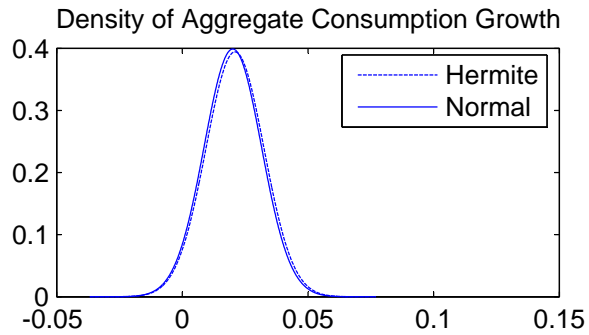Notes: This figure shows multivariate quantile-quantile (QQ) plots of log consumption growth and asset returns. Each panel plots the sample quantiles (on the y-axis) against the quantiles of a given distribution (on the x-axis) as well pointwise 5% and 95% bands. The top panel shows the QQ plot for the joint distribution of $\Delta c, r_s$ and r_f, i.e. the quantiles of the squared Mahalanobis distances against those of a χ^2_3 distribution. The bottom panel shows the QQ plot for the joint distribution of $\Delta c, r_s, r_f$ and 6 FF portfolios, i.e. the quantiles of the squared Mahalanobis distances against those of a χ^2_9 distribution. The squared Mahalanobis distance M_t for a p-dimensional multivariate distribution $\boldsymbol{x}_t$ with mean $\boldsymbol{\mu}_x$ and variance-covariance matrix $\boldsymbol{V}$ is defined as $M_t = (\boldsymbol{x}_t - \boldsymbol{\mu}_x)'\boldsymbol{V}^{-1}(\boldsymbol{x}_s - \boldsymbol{\mu}_x)$. Under the null hypothesis that $\Delta c, r_s$ and r_f are jointly normally distributed, M_t has a χ^2_p distribution.

Figure 4

Density of Aggregate Consumption Growth

Density of Assetholder Consumption Growth

Density of Dividend Growth

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

Figure 5

Density of Aggregate Consumption Growth

Density of Aggregate Consumption Growth

Density of Assetholder Consumption Growth

Density of Assetholder Consumption Growth

Density of Dividend Growth

Density of Dividend Growth

Notes: Plots of marginal densities for two Hermite parameter configurations that do not explain large pricing errors of the standard model.