Title
Fat tails and spurious estimation of consumption-based asset pricing models

Journal
Journal of Applied Econometrics, 32(6)

Authors
Toda, Alexis Akira
Walsh, Kieran James

Publication Date
2017-09-01

Peer reviewed

Fat Tails and Spurious Estimation of Consumption-Based Asset Pricing Models*

Alexis Akira Toda† Kieran James Walsh‡

November 17, 2016

Abstract

The standard generalized method of moments (GMM) estimation of Euler equations in heterogeneous-agent consumption-based asset pricing models is inconsistent under fat tails because the GMM criterion is asymptotically random. To illustrate this, we generate asset returns and consumption data from an incomplete-market dynamic general equilibrium model that is analytically solvable and exhibits power laws in consumption. Monte Carlo experiments suggest that the standard GMM estimation is inconsistent and susceptible to Type II errors (incorrect non-rejection of false models). Estimating an overidentified model by dividing agents into age cohorts appears to mitigate Type I and II errors.

Keywords: consumption-based CAPM, generalized method of moments, heterogeneous-agent model, power law

JEL codes: C58, D31, D52, D58, G12

1 Introduction

It is well-known that the representative-agent, consumption-based capital asset pricing model (CCAPM) of Lucas (1978) and Breeden (1979) requires a relative risk aversion coefficient on the order of 10 to 100 in order to explain the historical equity premium, at least in the basic, frictionless case with additively separable, constant relative risk aversion (CRRA) preferences.[1] To explain asset prices with a lower risk aversion parameter, many researchers have considered the possibility of market incompleteness and estimated and tested *heterogeneous-agent*

*We thank Bertille Antoine, Brendan Beare, Chris Carroll, Russell Davidson, Lynda Khalaf, Yixiao Sun, and seminar participants at Australian School of Business, Université Laval, UCSD, Yale, the 17th ICMAIF at the University of Crete, and the CIREQ Conference on Financial Econometrics for comments and feedback. We especially thank two anonymous referees for suggestions that significantly improved the paper.

†Department of Economics, University of California San Diego. 9500 Gilman Dr, La Jolla, CA 92130. Email: atoda@ucsd.edu

‡Darden School of Business, University of Virginia. 100 Darden Blvd, Charlottesville, VA 22903. Email: walshk@darden.virginia.edu

[1]See Grossman and Shiller (1981), Hansen and Singleton (1983), Mehra and Prescott (1985), Grossman et al. (1987), Kocherlakota (1997), and Savov (2011), among many others.

models[2] using household consumption data such as the Consumer Expenditure Survey (CEX).

In a recent paper (Toda and Walsh, 2015), we document that the cross-sectional distributions of U.S. household consumption and its growth rate exhibit fat tails. The power law exponent $\alpha > 0$ is approximately four both in the upper and lower tails.[3] If this is the case, the cross-sectional moments of consumption and its growth rate, $\mathrm{E}_t[c_{it}^\eta]$ and $\mathrm{E}_t[(c_{it}/c_{i,t-1})^\eta]$, are infinite when the moment order η exceeds the power law exponent α in absolute value. Such nonexistence of moments renders the generalized method of moments (GMM) estimation of aggregated household Euler equations inconsistent due to lack of identification: even if the model is correctly specified, nonexistent moments aid in zeroing the GMM criterion at untrue parameters (a Type I error). Furthermore, our bootstrap studies suggest that the fat tails in consumption mechanically set the pricing error to zero, even when the model is incorrect (a Type II error). As we show in Section 2.2, the problem is that when the moment conditions contain nonexistent cross-sectional moments, the criterion function becomes asymptotically random. The implication is that GMM estimation may find a spurious criterion minimum due to randomness rather than to the truth of the model.

As a remedy, Toda and Walsh (2015) propose an alternative estimation approach (age cohort GMM) that divides households into age groups in order to mitigate the fat tail problem. This approach is motivated by the finding in Battistin et al. (2009) that, *within age cohorts*, the empirical cross-sectional distribution of consumption is approximately lognormal, which is thin-tailed.[4] However, the analysis of Toda and Walsh (2015) is only suggestive since, with actual consumption data, we know neither the true data generating process nor whether the asset pricing model is a good description of reality.

In this paper, we conduct a Monte Carlo study using artificial asset returns and consumption data.[5] The goal is to assess the robustness (or non-robustness) of estimation/testing of heterogeneous-agent asset pricing models when the cross-sectional consumption distribution exhibits fat tails and the models may be true *or* false. Compared to the representative-agent setting, a simulation study of a heterogeneous-agent model is challenging for two reasons. First, solving a heterogeneous-agent asset pricing model is much more com-

[2]See Mankiw (1986), Constantinides and Duffie (1996), Heaton and Lucas (1996), Saito (1998), Krebs and Wilson (2004), Storesletten et al. (2007), Krueger and Lustig (2010), Guvenen (2009), and Toda (2015) for theoretical/numerical works and Brav et al. (2002), Cogley (2002), Vissing-Jørgensen (2002), Balduzzi and Yao (2007), Krueger et al. (2008), Kocherlakota and Pistaferri (2009), Basu et al. (2011), Constantinides and Ghosh (2014), and Semenov (2016) for empirical works. See Ludvigson (2013) for a review on testing CCAPM.

[3]A nonnegative random variable X is said to be *Paretian* (obey the power law in the upper tail) if $\Pr(X > x) = Ax^{-\alpha}(1 + o(1))$ as $x \to \infty$ for some $A, \alpha > 0$. It obeys the power law in the lower tail if $1/X$ is Paretian, so $\Pr(X < x) = Bx^{\beta}(1 + o(1))$ as $x \to 0$ for some $B, \beta > 0$. $\alpha, \beta > 0$ are called power law exponents. See Resnick (2008) for an authoritative textbook treatment of extreme value theory and Gabaix (2009) for a review of empirical power laws as well as some generative mechanisms.

[4]Battistin et al. (2009) document the lognormality in consumption using U.S. and U.K. data. Brzozowski et al. (2010) obtain similar results with Canadian data.

[5]Thus our approach is similar in spirit to Tauchen (1986), Kocherlakota (1990), and Hansen et al. (1996), who study the finite sample properties of the GMM estimator of representative-agent asset pricing models. Carroll (2001) uses simulated data to estimate the relative risk aversion from the log-linearized Euler equation. Alan et al. (2009) study the finite sample properties of a GMM estimator that is robust to measurement error.

plicated than solving a representative-agent one: heterogeneous-agent models rarely have closed-form solutions, and one must thus usually solve them numerically as in Krusell and Smith (1998). Solving even two agent, two asset infinite horizon general equilibrium models often entails substantial computational burden (Guvenen, 2009). Second, since our aim is to study the implications of fat tails for GMM estimation, the cross-sectional distribution of consumption must have fat tails. But, numerical techniques do not let us, with certainty, identify or characterize fat tails in heterogeneous-agent models.

The incomplete-market dynamic general equilibrium model of Toda (2014) overcomes both difficulties: it is analytically solvable and computationally tractable, and the model's cross-sectional consumption distribution obeys the power law in both the upper and lower tails with known power law exponents. Although in the literature there exist heterogeneous-agent models that are analytically solvable and exhibit fat tails, such as Benhabib et al. (2011, 2016), these models do not feature aggregate shocks and therefore cannot be applied to the study of asset prices. The Toda (2014) model, on the other hand, allows for an arbitrary Markov process for the aggregate shocks. Therefore, this model is well-suited as a laboratory for examining financial Euler equation estimation in the presence of fat tails in the cross-sectional distribution of consumption.

Using the incomplete-market general equilibrium model as our laboratory, we conduct two sets of experiments. First, we estimate the relative risk aversion coefficient both by standard GMM and by age cohort GMM using the simulated consumption and asset returns data from the model. We find that standard GMM over-rejects the correct risk aversion coefficient and that the GMM estimator has a large mean squared error. In part, this is because in many simulations, the GMM criterion has multiple troughs, one near the true parameter and another at a random location, which is frequently the global minimum. The risk aversion estimate is often well above 10 (in the moment nonexistence range) and, in these instances, associated with a zero equity premium pricing error. Thus, the fat tails sometimes aid in over-fitting, even though on average standard GMM over-rejects the correct model. Age cohort GMM, in contrast, provides more accurate risk aversion estimates and does not over-reject the true model/parameter.

Second, we repeat our analysis but with incorrect, random asset returns. Standard GMM quite often fails to reject the model even though it is false. In these cases of over-fitting, risk aversion estimates are upwardly biased high into the moment nonexistence range. Oddly, the histogram of equity premium pricing errors across simulations is bimodal, with spurious mass at zero. The age cohort method, on the other hand, removes the spurious pricing error peak at zero. While some of these findings seem odd, they closely mirror the empirical findings of Toda and Walsh (2015).

Our results are driven by the sample analogs of nonexistent moments and not by GMM per se: generalized empirical likelihood (GEL) estimation of the model mostly generates worse Type I and Type II errors (see Online Appendix). Our findings are also robust across econometric specifications. Our baseline (the "conditional" model) uses a single risky asset and one instrument (the lagged price-dividend ratio), but the results are similar when we exclude the instrument (the "exactly identified" model). Dropping the instrument and adding another risky asset (the "unconditional" model) somewhat mitigates the spurious pricing error mode but still generates excessive Type I and II errors relative to age cohort

GMM.

Finally, using a representative-agent asset pricing model as an example, we provide a simple explanation for the bimodal pricing error histograms and the Type II errors. The idea is that since the sample moment condition involves negative powers of consumption growth, when the power is a large negative value (corresponding to high risk aversion), the sample moment condition is dominated by the two largest terms in absolute value (corresponding to the two smallest consumption growth observations). Therefore as long as these two terms have opposite signs, regardless of whether the model is true or false, we can always set the sample moment to zero at some high risk aversion parameter. Indeed, estimating a representative-agent model with the incorrect asset returns (so the model is false by construction), we get bimodal pricing errors and Type II errors, driven by high risk aversion. This explanation offers a simple remedy: estimating an overidentified model (with consumption powers in each moment condition) is likely to mitigate the Type II errors since the spurious estimators differ across equations. The age cohort method is an example of this remedy.

Our paper is related to two strands of literature. The first is the literature on model estimation under fat tails. In an asset pricing setting, Kocherlakota (1997) tests the standard representative-agent CCAPM with fat-tailed pricing errors using subsampling. Here the fat tails appear in the time series, whereas in our analysis they appear in the cross-section of consumption. While he tests the model using actual data, our focus is the estimation from simulated data, for which we know by construction both the data generating process and whether the model is true or false.[6] Beaulieu et al. (2010) test the Fama-French multifactor CAPM with fat-tailed asset returns. Although not in a financial context, Geweke (2001) describes the limitations of the constant relative risk aversion (CRRA) utility function: expected utility may not exist when the distribution of consumption is fat-tailed. In the GARCH setting, when error moments become infinite between 2 and 4, the convergence rate of quasi-maximum likelihood (QML) estimation falls below $n^{1/2}$ (Berkes and Horváth, 2003), and the asymptotic distribution may be non-Gaussian and difficult to estimate (Hall and Yao, 2003). To address fat tail issues in the GARCH context, Hill (2015) and Hill and Prokhorov (2016) introduce, respectively, tail-trimmed QML and tail-trimmed GEL, each of which yields asymptotic normality and better finite sample properties relative to a variety of standard methods.

The second literature concerns the problems with estimating asset pricing models with model misspecification or unidentified parameters. Kan and Zhang (1999a,b) develop the asymptotic theory and conduct simulations of the two-pass and the GMM tests of linear factor models that contain factors that are uncorrelated with asset returns ("useless factors"). They find that when the model is misspecified, the presence of useless factors leads to Type II errors. Kan et al. (2013) and Gospodinov et al. (2014) develop a misspecification-robust inference and model selection method for the two-pass and GMM tests, respectively. Burnside (2016) considers the case in which factor loadings are unidentified and theoretically shows that the estimation results are sensitive to the way one normalizes the stochastic discount factor. More broadly, Andrews and Cheng (2012) study the asymptotic properties of extremum estimators when there is

[6]Kocherlakota (1990) compares various tests of the representative-agent CCAPM using simulated data.

weak identification in parts of the parameter space. Their leading example is ML estimation of the ARMA(1,1) model, which becomes unidentified as the true AR and MA coefficients approach one another. Simulations show that in this case estimate distributions are bimodal and standard tests over-reject correct nulls. Our study provides another practical example of poor identification leading to multi-humped estimate distributions and Type I errors.

Our paper is at the intersection of these literatures because in our setting fat tails lead to inconsistency and Type II errors. Specifications that average across Euler equations introduce nonexistent moments, which cause the GMM criterion to be asymptotically random. Complementary to the tail-trimming explored in Hill (2015) and Hill and Prokhorov (2016), our Monte Carlo experiments show that age cohort GMM, and overidentifying restrictions in general, yield improvements in mean squared error and test size/power.

2 Euler equation aggregations and inconsistency under fat tails

2.1 Literature on Euler equation aggregations[7]

Consider an economy populated by households with identical additive constant relative risk aversion (CRRA) preferences

$$\mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\gamma}}{1-\gamma},$$

where $\beta > 0$ is the discount factor, $\gamma > 0$ is the relative risk aversion coefficient, and c_t is consumption. Assuming interior solutions, the Euler equation

$$c_{it}^{-\gamma} = \mathrm{E}\left[\beta c_{i,t+1}^{-\gamma} R_{t+1} \,\middle|\, \mathcal{F}_{it}\right] \tag{2.1}$$

holds, where R_{t+1} is the gross return of any asset and $\mathcal{F}_{it}$ denotes the information set of household i at time t.

In order to estimate and test these Euler equations using micro consumption data, one must overcome two potential problems: measurement error in household-level consumption and panel shortness (individual households participate for only short periods of time). To handle these issues, the empirical literature on testing heterogeneous-agent asset pricing models "averages" across households to mitigate measurement error and create a long time series. This literature has provided several approaches to aggregating the Euler equations.

The first approach is to average the marginal rate of substitution as in Brav et al. (2002) and Cogley (2002), which are based on the theoretical model of Constantinides and Duffie (1996). Let $\mathcal{F}_t$ be the information set that contains only aggregate variables—in this example asset returns—and let E_t denote the expectation conditional on $\mathcal{F}_t$. Dividing (2.1) by $c_{it}^{-\gamma}$, conditioning on aggregate variables $\mathcal{F}_t$, and applying the law of iterated expectations, we obtain

$$1 = \mathrm{E}_t\left[\beta (c_{i,t+1}/c_{it})^{-\gamma} R_{t+1}\right] = \mathrm{E}_t\left[\beta\, \mathrm{E}_{t+1}[(c_{i,t+1}/c_{it})^{-\gamma}] R_{t+1}\right].$$

[7] This section draws heavily from the literature review of Toda and Walsh (2015) in order to make the paper self-contained.

Since this equation holds for any asset, subtracting the equation corresponding to the risk-free rate R_t^f and dividing by $\beta > 0$, we obtain

$$\mathrm{E}_t\left[m_{t+1}^{\mathrm{IMRS}}(R_{t+1} - R_t^f)\right] = 0,$$

where $m_{t+1}^{\mathrm{IMRS}} = \mathrm{E}_{t+1}[(c_{i,t+1}/c_{it})^{-\gamma}]$ is the $-\gamma$-th cross-sectional moment of consumption growth between time t and $t+1$. Therefore, up to a multiplicative constant (here β), m_{t+1}^{IMRS} is a valid stochastic discount factor (SDF), where IMRS stands for "intertemporal marginal rate of substitution". For estimation, we can use the sample analog

$$\widehat{m}_{t+1}^{\mathrm{IMRS}}(\gamma) = \frac{1}{I}\sum_{i=1}^{I}\left(\frac{c_{i,t+1}}{c_{it}}\right)^{-\gamma} \tag{2.2}$$

and minimize the GMM criterion

$$J_T(\gamma) = T\left(\frac{1}{T}\sum_{t=1}^{T}\widehat{m}_t^{\mathrm{IMRS}}(\gamma)(R_t^s - R_{t-1}^f)\right)^2, \tag{2.3}$$

where R_t^s is the stock return.

One issue with the IMRS SDF is that, since it is the cross-sectional average of the negative power of individual consumption growth, its value will be highly sensitive to the smallest consumption growth observation or measurement error. As a remedy, Balduzzi and Yao (2007) have suggested a more robust SDF by averaging the Euler equation (2.1) directly. Taking the expectation of (2.1) with respect to $\mathcal{F}_t$ and applying the law of iterated expectations, we obtain

$$\mathrm{E}_t[c_{it}^{-\gamma}] = \mathrm{E}_t[\beta c_{i,t+1}^{-\gamma}R_{t+1}] = \mathrm{E}_t\left[\beta\,\mathrm{E}_{t+1}[c_{i,t+1}^{-\gamma}]R_{t+1}\right].$$

Dividing both sides by $\mathrm{E}_t[c_{it}^{-\gamma}]$, we obtain

$$1 = \mathrm{E}_t\left[\beta\frac{\mathrm{E}_{t+1}[c_{i,t+1}^{-\gamma}]}{\mathrm{E}_t[c_{it}^{-\gamma}]}R_{t+1}\right].$$

By the same argument as above,

$$m_{t+1}^{\mathrm{MU}} = \frac{\mathrm{E}_{t+1}[c_{i,t+1}^{-\gamma}]}{\mathrm{E}_t[c_{it}^{-\gamma}]}$$

is also a valid stochastic discount factor up to a multiplicative constant, where MU stands for "marginal utility". For estimation, we can use the sample analog

$$\widehat{m}_{t+1}^{\mathrm{MU}}(\gamma) = \frac{\frac{1}{I}\sum_{i=1}^{I}c_{i,t+1}^{-\gamma}}{\frac{1}{I}\sum_{i=1}^{I}c_{it}^{-\gamma}}. \tag{2.4}$$

Balduzzi and Yao (2007) argue that the MU SDF is less susceptible to measurement error, because if the process for measurement error is i.i.d. across agents (but not necessarily over time), then the term corresponding to the measurement error will cancel out in the numerator and the denominator of m_{t+1}^{MU}.

As pointed out by Toda and Walsh (2015), the validity of the IMRS and MU stochastic discount factors relies on the existence of the cross-sectional moments $\mathrm{E}_t[(c_{it}/c_{i,t-1})^{-\gamma}]$ and $\mathrm{E}_t[c_{it}^{-\gamma}]$, respectively. However, the above studies do not explicitly discuss the presence or implications of fat tails in the cross-sectional distribution of consumption or consumption growth.

2.2 Inconsistency of GMM under fat tails

Why might fat tails in the consumption distribution create problems for GMM estimation? We can illustrate the problem in a very simple setting. Suppose that $\{x_t, y_t\}_{t\in\mathbb{Z}}$ is i.i.d., $\mathrm{E}[x_t^2] < \infty$, and $y_t = \theta_0 x_t + \epsilon_t$, where the error term ϵ_t is independent from x_t. Suppose the researcher believes that $\mathrm{E}[\epsilon_t] = 0$ and uses the moment condition

$$\mathrm{E}[(y_t - \theta x_t)z_t] = 0 \iff \theta = \theta_0$$

to estimate θ by GMM (in this case, method of moments), where $z_t = x_t$ is the regressor used as an instrument. Clearly the GMM (OLS) estimator is

$$\widehat{\theta}_T = \frac{T^{-1}\sum_{t=1}^T y_t z_t}{T^{-1}\sum_{t=1}^T x_t z_t} = \theta_0 + \frac{T^{-1}\sum_{t=1}^T x_t \epsilon_t}{T^{-1}\sum_{t=1}^T x_t^2},$$

where T is the sample size. If indeed $\mathrm{E}[\epsilon_t] = 0$, by the strong law of large numbers we have

$$\widehat{\theta}_T \xrightarrow{\text{a.s.}} \theta_0 + \frac{\mathrm{E}[x_t\epsilon_t]}{\mathrm{E}[x_t^2]} = \theta_0 + \frac{\mathrm{E}[x_t]\,\mathrm{E}[\epsilon_t]}{\mathrm{E}[x_t^2]} = \theta_0,$$

so $\widehat{\theta}_T$ is consistent.

Now suppose, in fact, that $|\epsilon_t|$ is Paretian with exponent $0 < \alpha < 1$. By Theorem 3 of Embrechts and Goldie (1980) (see also Cline (1986)), $x_t\epsilon_t$ also has a power law exponent α. Consequently, as is well-known (*e.g.*, Theorem 9.34 and Problem 9.10 in Breiman (1968), Theorem 3.7.2 and Exercise 3.7.2 in Durrett (2010)), it follows that

$$T^{-1/\alpha}\sum_{t=1}^T x_t\epsilon_t \xrightarrow{\text{d}} Y,$$

where Y is a nondegenerate distribution (a suitably normalized Lévy α-stable distribution). Therefore

$$\widehat{\theta}_T = \theta_0 + T^{1/\alpha-1}\frac{T^{-1/\alpha}\sum_{t=1}^T x_t\epsilon_t}{T^{-1}\sum_{t=1}^T x_t^2} \sim \theta_0 + T^{1/\alpha-1}\frac{Y}{\mathrm{E}[x_t^2]},$$

and since $1/\alpha - 1 > 0$, the GMM estimator $\widehat{\theta}_T$ diverges and hence is inconsistent. ($T^{1-1/\alpha}\widehat{\theta}_T$ converges in distribution to $Y/\mathrm{E}[x_t^2]$.) The problem is that the GMM criterion

$$\begin{aligned}\left(\frac{1}{T}\sum_{t=1}^T (y_t - \theta x_t)x_t\right)^2 &= \left(T^{1/\alpha-1}T^{-1/\alpha}\sum_{t=1}^T x_t\epsilon_t - (\theta-\theta_0)\frac{1}{T}\sum_{t=1}^T x_t^2\right)^2 \\ &\sim (T^{1/\alpha-1}Y - (\theta-\theta_0)\,\mathrm{E}[x_t^2])^2\end{aligned}$$

diverges almost surely as $T \to \infty$, or once rescaled is random asymptotically. Although we have maintained an i.i.d. assumption for simplicity, we obtain the same conclusion in the non-i.i.d. case by using the results of Davis and Hsing (1995).

The same issue applies to the IMRS stochastic discount factor (2.2). Suppose, for simplicity, that aggregate consumption growth $G_{t+1} := C_{t+1}/C_t$ is i.i.d. over time, and that the growth rate of individual consumption relative to the aggregate consumption, $g_{i,t+1} := \frac{c_{i,t+1}/C_{t+1}}{c_{it}/C_t}$, is also i.i.d. over time and across individuals. Furthermore, assume that $1/g_{i,t+1}$ has a power law with exponent $\alpha > 0$. (In the data, $\alpha \approx 4$.) If $\gamma > \alpha$, since $g_{i,t+1}^{-\gamma}$ has a power law exponent $\alpha/\gamma < 1$, by the same argument as above, we have

$$\begin{aligned}\widehat{m}_{t+1}^{\text{IMRS}}(\gamma) = \frac{1}{I}\sum_{i=1}^{I}\left(\frac{c_{i,t+1}}{c_{it}}\right)^{-\gamma} &= \frac{1}{I}G_{t+1}^{-\gamma}\sum_{i=1}^{I} g_{i,t+1}^{-\gamma} \\ &\sim I^{\gamma/\alpha-1}G_{t+1}^{-\gamma}Y_{t+1}(\gamma),\end{aligned}$$

where $Y_{t+1}(\gamma)$ has a nondegenerate distribution that depends on γ. Since by assumption $g_{i,t+1}$ is i.i.d. over time and individuals, $\{Y_{t+1}(\gamma)\}$ is i.i.d., and is a suitably normalized stable distribution with index $\alpha/\gamma < 1$. Letting $X_t = R_t^s - R_{t-1}^f$ be the excess return, the expression inside the parenthesis of the IMRS GMM criterion (2.3) is

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T}\widehat{m}_t^{\text{IMRS}}(\gamma)X_t &\sim T^{\gamma/\alpha-1}I^{\gamma/\alpha-1}T^{-\gamma/\alpha}\sum_{t=1}^{T}G_t^{-\gamma}Y_t(\gamma)X_t \\ &\sim T^{\gamma/\alpha-1}I^{\gamma/\alpha-1}Z(\gamma),\end{aligned}$$

where again $Z(\gamma)$ is a suitably normalized stable distribution. Thus the GMM estimator will asymptotically behave as the minimizer or $Z(\gamma)^2$, which is a random function, and hence the GMM estimator is inconsistent. A similar argument holds for the MU SDF as well.

Given these theoretical results, we can expect that the standard GMM estimation of heterogeneous-agent asset pricing models will have poor properties. However, in finite samples would the estimator be biased upwards or downwards? Would standard tests lead to over or under rejections? It is difficult to answer these questions with actual data since we know neither the true data generating process nor whether the model is true or false. Therefore we resort to a Monte Carlo study using simulated data.

3 Simulating an economy

In this section we generate asset returns and consumption data from an incomplete-market dynamic general equilibrium model that admits a closed-form solution. Because the model is highly tractable and the cross-sectional consumption distribution obeys the power law in both tails, we can create an artificial economy with a consumption distribution that has fat tails with known power law exponents and then use it as a laboratory for studying the properties of the MU stochastic discount factor, which would be valid in this setting if not for fat fails.

3.1 Model

We present a minimal model to simulate an economy with a fat-tailed consumption distribution.

3.1.1 Settings

We consider a heterogeneous-agent, consumption-based asset pricing model similar to Constantinides and Duffie (1996). Time is indexed by $t = 0, 1, \dots$ and agents are indexed by $i \in I = \{1, \dots, I\}$. As in Blanchard (1985), between consecutive periods each agent dies at constant probability $0 < \delta < 1$ independently across agents and time, and is replaced by a newborn agent. This overlapping generation feature is necessary in order to obtain a non-degenerate cross-sectional distribution. Agents have identical standard additive CRRA preferences

$$\mathrm{E}_0 \sum_{t=0}^{\infty} (\beta(1-\delta))^t \frac{c_{it}^{1-\gamma}}{1-\gamma},$$

where $\beta > 0$ is the discount factor, $(1-\delta)^t$ is the probability to survive up to time t, $\gamma > 0$ is the relative risk aversion coefficient, and c_{it} is agent i's consumption.

There are three assets, a claim to aggregate dividends (dividend claim), a claim to aggregate consumption (consumption claim), and a one-period risk-free bond, all in zero net supply.[8] Let C_t, D_t be aggregate consumption and dividends. The aggregate endowment is denoted by Y_t. Let

$$x_t = (\log(Y_t/Y_{t-1}), \log(D_t/D_{t-1}))'$$

be the vector of log aggregate endowment and dividend growth. Since it is a pure exchange economy, aggregate consumption C_t equals the aggregate endowment Y_t by market clearing. We assume that x_t obeys a VAR(1) process

$$x_t = (I - A)g + Ax_{t-1} + u_t, \quad u_t \sim N(0, \Sigma), \tag{3.1}$$

where A is a 2×2 matrix with all eigenvalues less than 1 in absolute value, $g = (g_c, g_d)'$ is the unconditional mean of log aggregate consumption and dividend growth, and u_t is an error term that is i.i.d. over time.

Assume that for surviving agents, log individual endowment growth equals aggregate endowment growth plus an uninsurable idiosyncratic shock:

$$\log \frac{y_{it}}{y_{i,t-1}} = \log \frac{Y_t}{Y_{t-1}} + \varepsilon_{it}, \quad \varepsilon_{it} \sim N(-\sigma^2/2, \sigma^2), \tag{3.2}$$

where $\sigma > 0$ is the idiosyncratic volatility. For simplicity, the idiosyncratic shock ε_{it} is assumed to be i.i.d. across individual and time. Note that since $\varepsilon_{it} \sim N(-\sigma^2/2, \sigma^2)$, we have $\mathrm{E}[\mathrm{e}^{\varepsilon_{it}}] = 1$. ε_{it} determines inequality over the life cycle.

For agents that are reborn, the initial endowment equals the aggregate endowment times a lognormal idiosyncratic shock:

$$\log y_{it} = \log Y_t + \eta_{it}, \quad \eta_{it} \sim N(-\sigma_0^2/2, \sigma_0^2),$$

where η_{it} determines the innate inequality.

This economy is tractable enough so that we can compute all asset prices in closed-form. See Appendix A for details.

[8] The zero net supply assumption is innocuous since if the assets are in positive net supply, by giving shares to individuals proportional to their income (at $t = 0$ or at birth), we can construct an equilibrium with no trade as in Constantinides and Duffie (1996).

3.1.2 Cross-sectional distribution

Next, we characterize the consumption distribution. Invoking the equilibrium condition $c_{it} = y_{it}$ and $C_t = Y_t$ in (3.2), the logarithm of individual consumption relative to aggregate consumption satisfies

$$\log \frac{c_{it}}{C_t} = \log \frac{c_{i,t-1}}{C_{t-1}} + \varepsilon_{it}.$$

Since $\varepsilon_{it} \sim N(-\sigma^2/2, \sigma^2)$, the log relative consumption is a random walk with a drift $\mu = -\sigma^2/2$ and instantaneous variance σ^2. Since endowment at birth is lognormal, the cross-sectional distribution within an age cohort is also lognormal. The log variance of a cohort with age a is $\sigma_0^2 + \sigma^2 a$, which increases linearly with age.

Since agents die at constant probability $0 < \delta < 1$ between each period and are reborn, the age distribution is geometric with mean $1/\delta$. Since the cross-sectional consumption distribution within each age cohort is lognormal and the log variance increases linearly with age, the entire cross-sectional log consumption distribution is a normal mixture. Under general settings, Toda (2014) shows that in the continuous-time limit, the shape of the cross-sectional distribution of consumption (relative to when born) converges to the *double Pareto* distribution (Reed, 2001), which is a distribution with two Pareto tails. The density function is

$$f(x) = \begin{cases} \frac{\alpha_1\alpha_2}{\alpha_1+\alpha_2} x^{-\alpha_1-1}, & (x \geq 1) \\ \frac{\alpha_1\alpha_2}{\alpha_1+\alpha_2} x^{\alpha_2-1}, & (x \leq 1) \end{cases}$$

where α_1, α_2 are the power law exponents of the upper and lower tails. According to Theorem 16 of Toda (2014), $-\alpha_1$ and α_2 are solutions to the quadratic equation

$$\frac{\sigma^2}{2}\zeta^2 - \mu\zeta - \delta = 0.$$

Substituting $\mu = -\sigma^2/2$ and solving the equation, the power law exponents are

$$\alpha_1, \alpha_2 = \frac{1}{2}\left(\sqrt{1 + \frac{8\delta}{\sigma^2}} \pm 1\right), \tag{3.3}$$

where $\sigma > 0$ is the idiosyncratic volatility. In this case the cross-sectional moment of consumption $\mathrm{E}_t[c_{it}^\eta]$ is finite if and only if $-\alpha_2 < \eta < \alpha_1$. When δ is large compared to σ^2, then we have $\alpha_1, \alpha_2 \approx \sqrt{2\delta}/\sigma \pm 1/2$, so the average of the power law exponents is about $\sqrt{2\delta}/\sigma$.

Since the individual endowment is lognormally distributed when agents are born, the actual cross-sectional consumption distribution will be the product of lognormal and double Pareto distributions, which is known as the *double Pareto-lognormal* (Reed, 2003). This distribution is determined by four parameters, the mean and variance of the lognormal component and the two power law exponents of the double Pareto component. In our case, the variance parameter is σ_0 and the power law exponents are α_1, α_2 in (3.3).

3.2 Calibration

We calibrate an economy at the annual frequency. We assume no discounting, so $\beta = 1$. The death probability is $\delta = 1/30$, which implies an average lifespan of 30 years. As in Toda (2014), "death" in this model should not be taken literally and instead interpreted as the arrival of a major life event such as personal bankruptcy, retirement, divorce, death, etc. Under this interpretation, choosing an average of 30 years seems quite natural. The effective discount factor is then $\tilde{\beta} = \beta(1-\delta) = 0.967$, which is very close to values used in the literature. The relative risk aversion coefficient is $\gamma = 7$, which is arguably a little high but still lower than values used in many macro-finance papers.

For the dynamics of log consumption/dividend growth, we obtain the 1889–2009 real per capital consumption and real dividend from Robert Shiller's website[9] and estimate the VAR(1) process in (3.1) by ordinary least squares (OLS). The result is

$$\widehat{g} = \begin{bmatrix} .0203 \\ .0108 \end{bmatrix}, \quad \widehat{A} = \begin{bmatrix} -.0767 & .0119 \\ .8011 & .0592 \end{bmatrix}, \quad \widehat{\Sigma} = \begin{bmatrix} .0012 & .0015 \\ .0015 & .0125 \end{bmatrix}.$$

According to (3.3), the power law exponents are around $\sqrt{2\delta}/\sigma$. Since the estimate in Toda and Walsh (2015) is 4 in the U.S., we set the idiosyncratic volatility $\sigma = 0.0645$ to match the power law exponents. Deaton and Paxson (1994) find that the U.S. cross-sectional log variance within age cohorts increases almost linearly with age (which is consistent with our model), and the rate is .0069 per year. This value translates to an idiosyncratic volatility of $\sqrt{\frac{2}{3}.0069} = 0.0678$,[10] which is similar to our number (.0645). Finally, we assume that individual consumption is observed with a measurement error, with log standard deviation $\sigma_\epsilon = 0.1$ (10%).[11]

We can compute the price-dividend/consumption ratios, asset returns, and the risk-free rate by (A.3), (A.5), and (A.6) in Appendix A. With the above parameter values, the average price-dividend ratio (computed by integrating (A.3) with respect to the stationary distribution) is 32.8 (dividend yield 3.05%), average stock market return and volatility are 5.10% and 14.1%, and the average risk-free rate and volatility are 2.99% and 1.85%, which are of the same order of magnitude as in U.S. data.[12] The correlation between the aggregate consumption growth and the consumption and dividend claims are 0.94 and 0.60,

[9] `http://www.econ.yale.edu/~shiller/data.htm`

[10] The factor $\frac{2}{3}$ comes from the Grossman et al. (1987) adjustment for time-aggregated data, which is necessary because the power law exponents are computed using the continuous-time approximation.

[11] We experimented with various standard deviations for measurement error (including no measurement error), and the results were similar. The standard deviation of $\sigma_\epsilon = 0.1$ is taken from the simulation in Balduzzi and Yao (2007).

[12] According to the Shiller 1889–2009 data, the historical numbers are 7.69% for stock returns (18.4% volatility), 1.97% for the risk-free rate (5.80% volatility), and 4.29% for the dividend yield. Since in our model the idiosyncratic shock in consumption growth is i.i.d. across individual and time, the idiosyncratic shock does not affect the equity premium (though it lowers the risk-free rate), as shown by Krueger and Lustig (2010). It is not difficult to obtain a larger equity premium within heterogeneous-agent models by introducing either stochastic idiosyncratic volatility (Storesletten et al., 2007), multiple sectors and production (Toda, 2015), or rare disasters (Schmidt, 2015), but we stick to i.i.d. idiosyncratic shocks since our purpose is to simulate a simple economy with fat-tailed consumption data and reasonable returns, not to perform detailed and highly realistic calibration that resolves many asset pricing puzzles.

respectively, which are relatively high. The power law exponents for consumption computed by (3.3) are $\alpha_1 = 4.53$ for the upper tail and $\alpha_2 = 3.53$ for the lower tail.

3.3 Simulation

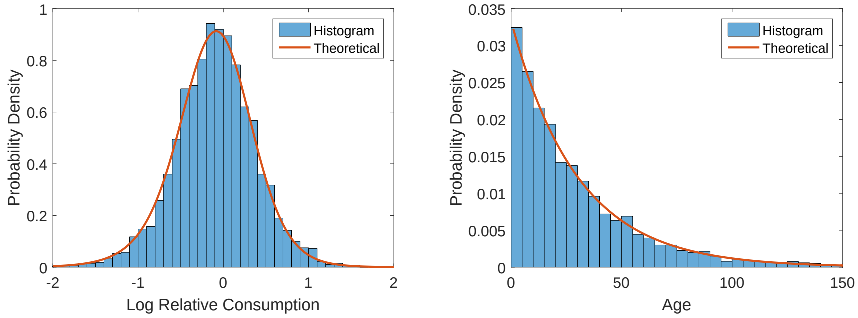
We simulate the economy with 10,000 Monte Carlo replications, each run consisting of either $T = 100$, 300, or 500 years and $I = 4000$ households at any given time.[13] The specific procedure is as follows. First, to create the panel of ages, we generate $I \times T$ Bernoulli variables with death probability $0 < \delta < 1$. Second, we set initial aggregate consumption $C_0 = 1$ and generate T aggregate shocks $\{x_t\}_{t=1}^T$, $I \times T$ idiosyncratic endowment growth shocks $\{(\varepsilon_{it})_{i\in I}\}_{t=1}^T$, $I \times T$ endowment level shocks at birth $\{(\eta_{it})_{i\in I}\}_{t=0}^{T-1}$, and compute the consumption path of each household denoted by $\{c_{it}\}$ as well as the stock return and the risk-free rate using (A.5) and (A.6). Finally, we multiply c_{it} by the "measurement error" e^ϵ, where $\epsilon \sim N(-\sigma_\epsilon^2/2, \sigma_\epsilon^2)$, again i.i.d. across agents and time. In this way we obtain a sequence of asset returns $\left\{(R_{t+1}^c, R_{t+1}^d, R_t^f)\right\}_{t=0}^{T-1}$ and an $I \times T$ panel of observed consumption and age.

Because the measurement error is lognormal, the cross-sectional (observed) consumption distribution for large enough time periods becomes approximately the product of double Pareto-lognormal and lognormal distributions, which is again double Pareto-lognormal. One may calculate the power law exponents α_1, α_2 either theoretically using (3.3) or numerically by estimating them by maximum likelihood using the log observed consumption distribution. (In our simulation they are almost the same number, as they should be.) We find that the shape of the cross-sectional distribution typically converges to a steady state after $10/\delta$ periods (10 times the average lifespan of households). In practice, we generate data for $b + T$ periods and discard the first b observations as burn in, with $b = \lfloor 10/\delta \rfloor$.

Figure 3.1 shows the histograms of log relative consumption and age at $t = 1$ for one simulation. Since the burn in period is $10/\delta = 300$, this is actually the 301st observation from the simulated data. The solid lines show the theoretical densities. For log consumption, the density is the convolution of the normal $N(-\tau^2/2, \tau^2)$ with $\tau^2 = \sigma_0^2 + \sigma_\epsilon^2$ (coming from idiosyncratic shock at birth and log measurement error) and the logarithm of double Pareto with exponents α_1, α_2 (which is known as *Laplace* (Kotz et al., 2001)). The resulting distribution is known as *normal-Laplace*, which is the logarithm of the double Pareto-lognormal and has a known closed-form density function (Reed and Jorgensen, 2004). Since the birth/death probability is constant, the theoretical age distribution is geometric (exponential). According to Figure 3.1, the theoretical densities closely track the histograms, so the continuous-time approximation is very good.

Although the histogram of log consumption is bell-shaped and may appear to be normal, actually it is far from normal. First, it is asymmetric because the two power law exponents $\alpha_1 = 4.53$ and $\alpha_2 = 3.53$ are distinct (the lower tail is fatter). Second, since consumption has power law tails, log consumption has exponential tails, which are fatter than those of the normal distribution. To see

[13] In the empirical analysis of Toda and Walsh (2015), there are about 300 time periods and 4,000 households.

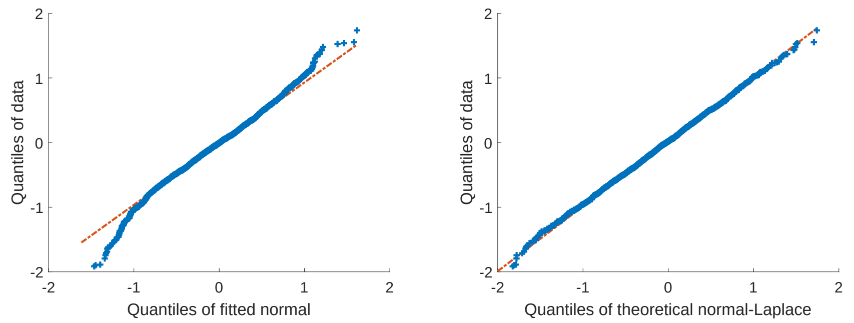


(a) Log relative consumption.

(b) Age.

Figure 3.1: Histograms of cross-sectional distributions at $t = 1$.

this graphically, Figure 3.2 shows the QQ (quantile-quantile) plot of log relative consumption against the normal distribution (fitted by maximum likelihood) and the normal-Laplace distribution (with the theoretical parameters). If the statistical model fits well to data, the QQ plot should show a 45 degree line. According to the result with the normal distribution (Figure 3.2a), the points deviate from the 45 degree line in the tails, which suggests that log consumption has much fatter tails than normal. On the other hand, the result with the normal-Laplace distribution (Figure 3.2b) shows a straight line, so the simulated data is close to the theoretical distribution.



(a) Normal.

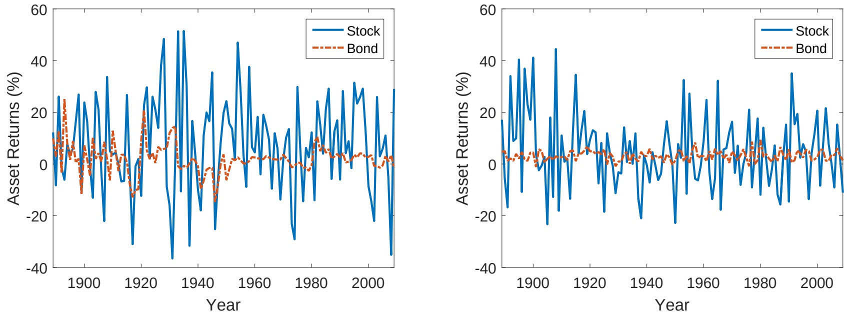
(b) Normal-Laplace.

Figure 3.2: QQ plot of log relative consumption at $t = 1$.

Figures 3.3 shows the actual (1889–2009) and simulated (first 121 years) asset returns, which show similar patterns.

4 Monte Carlo study

In our simulated data, we know that there is a power law in consumption, and we know that if not for this reason, the MU SDF would give us consistent estimates of γ, using simulated data. The question then is, how does the MU SDF behave in the presence of the power law? Since the true relative risk aversion coefficient ($\gamma = 7$) exceeds the power law exponent (4) and hence the



(a) Asset returns (actual data). (b) Asset returns (simulated data).

Figure 3.3: Asset returns from actual and simulated data.

cross-sectional moment of consumption does not exist at the true γ, we expect that the MU SDF will perform poorly because the large sample limit of the GMM criterion is random.

We consider the possibility of both a Type I error (incorrect rejection of a true null) and a Type II error (incorrect non-rejection of a false null). Two sorts of Type I errors are possible in our setting. First, the nonexistence of cross-sectional moments could prevent the true model from explaining the equity premium. Indeed, according to χ^2 tests of overidentifying restrictions, standard GMM over-rejects the true model. Second, inconsistency could lead us to find excessively high γ estimates and reject lower but correct values. This is what we find. Type II errors may arise precisely because the power law behavior lets us zero the pricing error at spuriously high γ estimates. Often, we fail to reject the model even when the asset returns are completely random.

4.1 GMM estimation

4.1.1 Standard GMM

The standard GMM proceeds as follows. Let

$$g_T(\gamma) = \frac{1}{T}\sum_{t=1}^{T} \widehat{m}_t^{\mathrm{MU}}(\gamma)(R_t^s - R_{t-1}^f) \otimes z_{t-1}$$

be the sample average of the pricing errors for the equity premium, where T is the number of time periods, $\widehat{m}_t^{\mathrm{MU}}(\gamma)$ is the MU stochastic discount factor in (2.4), R_t^s is the model-generated asset returns, R_{t-1}^f is the risk-free rate, and z_{t-1} is the vector of instruments. As described in the introduction, we consider three different specifications for R_t^s and z_{t-1}. For the "exactly identified" model, the only asset is the dividend claim ($R_t^s = R_t^d$), and there are no instruments. Instruments are not necessary for estimation but are necessary for tests of overidentifying restrictions if there is only one asset. Therefore, we also consider the "conditional" model. In this case, the dividend claim is still the only asset, but we use two instruments, the constant 1 and the normalized price-dividend ratio defined to be P_{t-1}/D_{t-1} divided by its sample mean. As the exactly identified and conditional models yield similar results, we focus on the latter, which allows for more tests. The third specification is the "unconditional" model, which has two assets, a claim on dividends and consumption ($R_t^s = (R_t^c, R_t^d)$), but no instruments. See Online Appendix for a comparison of all three specifications.

Letting W be the weighting matrix (we choose the identity matrix for the first stage estimation), the GMM estimator of the relative risk aversion coefficient γ and the mean squared pricing error are defined by

$$\widehat{\gamma} = \arg\min_{\gamma} T g_T(\gamma)' W g_T(\gamma),$$

$$e = \sqrt{\|g_T(\widehat{\gamma})\|^2 / K} = \|g_T(\widehat{\gamma})\| / \sqrt{K},$$

where K is the number of equations in GMM.[14] Since $\widehat{m}_t^{\mathrm{MU}}(\gamma)$ and z_{t-1} are numbers close to 1, the mean squared pricing error e has the same order of magnitude as the equity premium. This definition makes the comparison across different models intuitive, unlike the minimized GMM criterion which tends to be larger for overidentified models. Note, however, that since the first stage weighting matrix is the identity matrix, the mean squared error is just a monotonic transformation of the minimized GMM criterion. The calculation of standard errors and test statistics are explained in Online Appendix.[15]

In addition to standard GMM using the identity matrix as the weighting matrix, we also consider the generalized empirical likelihood (GEL) approach of Kitamura and Stutzer (1997) since GEL estimators are known to have smaller bias (Newey and Smith, 2004). Although there are many variants of GEL (see Kitamura (2007) for a review), the one that uses the Kullback-Leibler information as in Kitamura and Stutzer (1997) is particularly convenient because the dual optimization problem is unconstrained and low dimensional.

4.1.2 Age cohort GMM

As discussed in Section 2.2, the standard GMM is inconsistent when the consumption distribution has fat tails. To mitigate this issue, Toda and Walsh (2015) propose "age cohort GMM". Since the Euler equation aggregation in Section 2.1 that gave us the SDFs also works within a particular age cohort, and since the cross-sectional distribution of consumption is lognormal within age cohorts according to the model in Section 3, we can estimate an overidentified model by dividing agents into age groups. For example, divide the agents into H age groups according to the $100h/H$ percentile of the age distribution ($h = 1, \dots, H$), and call these groups $I_{t,1}, \dots, I_{t,H}$. We can form the MU SDF for cohort h by

$$\widehat{m}_{t,h}^{\mathrm{MU}}(\gamma) = \frac{\frac{1}{|I_{t,h}|}\sum_{i \in I_{t,h}} c_{it}^{-\gamma}}{\frac{1}{|I_{t-1,h}|}\sum_{i \in I_{t-1,h}} c_{i,t-1}^{-\gamma}},$$

where $|I_{t,h}|$ is the number of households in group $I_{t,h}$. One caveat is that since an agent with age a at time $t-1$ will have age $a+1$ at time t (if alive) and since

[14] In implementing the minimization over γ, to avoid local minima that are not the global minimum, we first perform a grid search over $\gamma = 0, 1, 2, \dots, 20$ and then use the minimizer as the initial value for the `fmincon` command in Matlab (with constraint $\gamma \geq 0$). We supply the analytical gradients to speed up the minimization.

[15] We also considered the efficient second stage estimation using the optimal weighting matrix, but we focus on the first stage because we find that the second stage estimator is biased, as reported in Altonji and Segal (1996) (linear model) and Clark (1996) (nonlinear model), and the bias and the standard errors are larger than in the first stage. Cochrane (2005) also recommends the first stage estimation for asset pricing models.

each Euler equation is agent specific, the age cutoffs for the numerator must be $+1$ of those of the denominator.

Let $\widehat{\mathbf{m}}_t^{\mathrm{MU}}(\gamma) = (\widehat{m}_{t,1}^{\mathrm{MU}}(\gamma), \dots, \widehat{m}_{t,H}^{\mathrm{MU}}(\gamma))'$ be the vector of SDFs and

$$G_T(\gamma) = \frac{1}{T}\sum_{t=1}^{T} \widehat{\mathbf{m}}_t^{\mathrm{MU}}(\gamma) \otimes (R_t^s - R_{t-1}^f) \otimes z_{t-1}$$

be the vector of pricing errors. Letting W be the weighting matrix, the first stage GMM estimator of γ and the mean squared pricing error are

$$\begin{aligned}
\widehat{\gamma} &= \arg\min_{\gamma} T G_T(\gamma)' W G_T(\gamma), \\
e &= \sqrt{\|G_T(\widehat{\gamma})\|^2 / (KH)} = \|G_T(\widehat{\gamma})\| / \sqrt{KH},
\end{aligned}$$

where K is the number of equations in each cohort and H is the number of cohorts. Below, we choose $H = 5$ (five age cohorts) and set the weighting matrix to the identity matrix.

4.1.3 Representative-agent GMM

Finally, as a robustness check, we also estimate γ from the representative-agent model (RA), which turns out to be valid for this particular example.[16] To see this, dividing both sides of the first-order condition (A.1) by $c_{it}^{-\gamma} P_t^d$, we obtain

$$1 = \tilde{\beta}\, \mathrm{E}_t \left[(c_{i,t+1}/c_{it})^{-\gamma} R_{t+1}^d \right],$$

where $R_{t+1}^d = (P_{t+1}^d + D_{t+1})/P_t^d$ is the dividend claim return. Since by (3.2) log individual consumption growth is equal to log aggregate consumption growth plus the idiosyncratic shock, it follows that

$$1 = \tilde{\beta} \mathrm{e}^{\frac{1}{2}\gamma(\gamma+1)\sigma^2}\, \mathrm{E}_t \left[(C_{t+1}/C_t)^{-\gamma} R_{t+1}^d \right].$$

The same equation holds for the consumption claim and the risk-free rate. Taking the difference and dividing by $\tilde{\beta} \mathrm{e}^{\frac{1}{2}\gamma(\gamma+1)\sigma^2}$, we obtain the moment condition

$$\mathrm{E}_t \left[(C_{t+1}/C_t)^{-\gamma} (R_{t+1}^d - R_t^f) \right] = 0.$$

Therefore up to a multiplicative constant, $m_{t+1}^{\mathrm{RA}}(\gamma) = (C_{t+1}/C_t)^{-\gamma}$ is also a valid stochastic discount factor. The GMM estimation of this representative-agent model is completely analogous.

[16] One could, in principle, also estimate the IMRS SDF because in our particular model consumption growth does not have fat tails and the measurement error is i.i.d. across agents and over time. We do not perform this exercise because (i) the IMRS SDF is more susceptible to measurement error issues in general (Balduzzi and Yao, 2007), (ii) empirical evidence suggests that the household consumption growth also has fat tails (Toda, 2016), and (iii) the point of this GMM exercise is to see what happens if we apply standard inferences when there are fat tail issues, not to carry out the most reasonable inference for this particular problem. By slightly changing the model (say by introducing time-varying idiosyncratic risk/measurement error or fat-tailed consumption growth) it is not difficult to make the IMRS SDF invalid.

4.2 Type I error

To study the possibility of a Type I error, we use the model-simulated consumption, stock, and bond return data and estimate each model by GMM. From top to bottom, Figure 4.1 shows the conditional model results for standard GMM, age cohort GMM, and representative-agent GMM with sample size $T = 500$.[17] The left panels show the scatter plot of simulated γ estimates and normalized mean squared pricing errors for 10,000 simulations. The right panels show the histogram of the pricing errors.

According to Figure 4.1a, across simulations there is an inverse relationship between the MU γ estimate and the pricing error. When the MU model almost exactly zeroes the pricing error, the γ estimate is often well above both the start of the moment nonexistence range, > 4, and the true coefficient, 7.

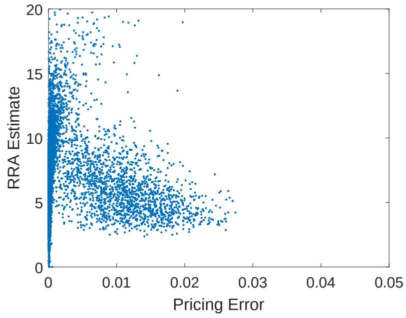
However, splitting households into age groups and performing the age cohort GMM, we no longer see this pattern: the large γ estimates corresponding to the zero pricing errors in Figure 4.1a have disappeared in the age cohort GMM of Figure 4.1c. Indeed, according to the histogram of the pricing errors in Figures 4.1b and 4.1d, there is much less mass around zero with the age cohort method. And, according to the scatter plots, this mass is the result of upwardly biased estimates in the nonexistence range. As we see in the scatter plot in Figure 4.1e, the RA γ estimates seem to be unbiased compared to the age cohort GMM, although they have larger standard errors because the representative-agent model exploits fewer moment restrictions. Also, the pricing errors are almost negligible. (Note that the scale of the horizontal axis is 10^{-3}.)

Figure 4.2 is the same as 4.1 but with the unconditional specification, which has consumption and dividend claims but no instruments. While the patterns are similar with respect to RA and age cohort GMM, standard GMM is somewhat improved: there is less pricing error mass at zero corresponding to upwardly biased estimates. As we discuss in Section 4.4, the improvements from the age cohort and unconditional specifications suggest overidentification mitigates the adverse impact of fat tails on standard GMM.

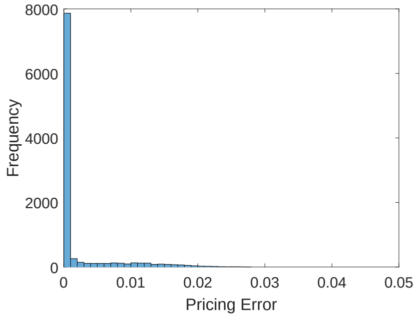
Table 4.1 show the bias (the average of $\widehat{\gamma} - \gamma$ across simulations), mean standard error truncated at 100 (to avoid excessively large numbers that appear in the standard and representative-agent GMM but not age cohort), mean absolute error (MAE, the average of $|\widehat{\gamma} - \gamma|$), and root mean squared error (RMSE, square root of the average of $|\widehat{\gamma} - \gamma|^2$) of each model/specification combination. For both the conditional and unconditional model, the age cohort GMM is the most biased but has the best finite sample properties in terms of standard error, mean absolute error, and root mean squared error. Using the unconditional model improves standard errors, MAEs, and RMSEs but worsens the bias for standard GMM (while lessening it with the other models).

Table 4.2 shows the Type I error probabilities, corresponding to a significance level of .05. For $T > 100$, in the standard GMM columns we see the manifestation of the high $\widehat{\gamma}$, low pricing error combinations in Figures 4.1a and 4.2a. With both the conditional and unconditional specifications, standard GMM over-rejects the true null ($\gamma = 7$), with sizes ranging from .075 to .092. In contrast, for $T > 100$ age cohort and RA sizes range from .040 to .052. χ^2 tests of overidentifying restrictions show that standard GMM also over-rejects

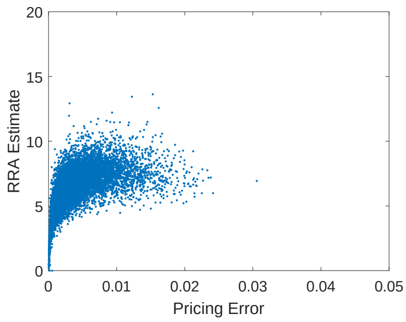
[17] We have also run the conditional model with $T = 1000$, which produced better finite sample properties but very similar figures.



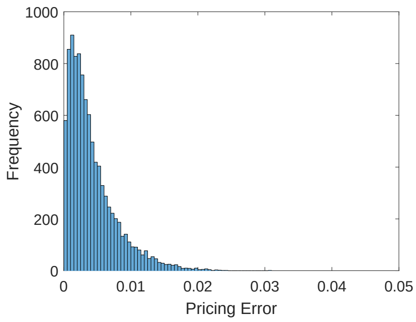
(a) Standard GMM



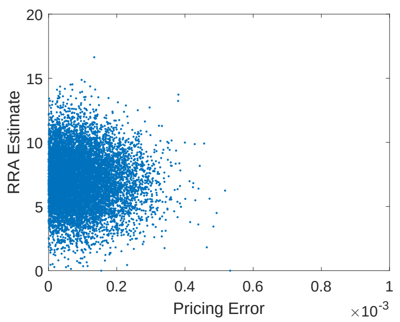
(b) Standard GMM



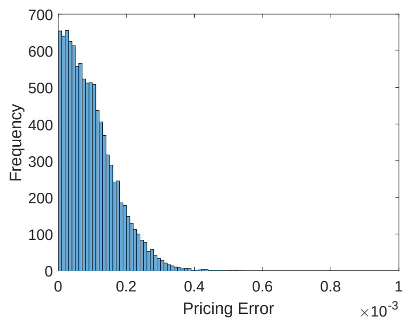
(c) Age cohort GMM



(d) Age cohort GMM



(e) Representative-agent GMM

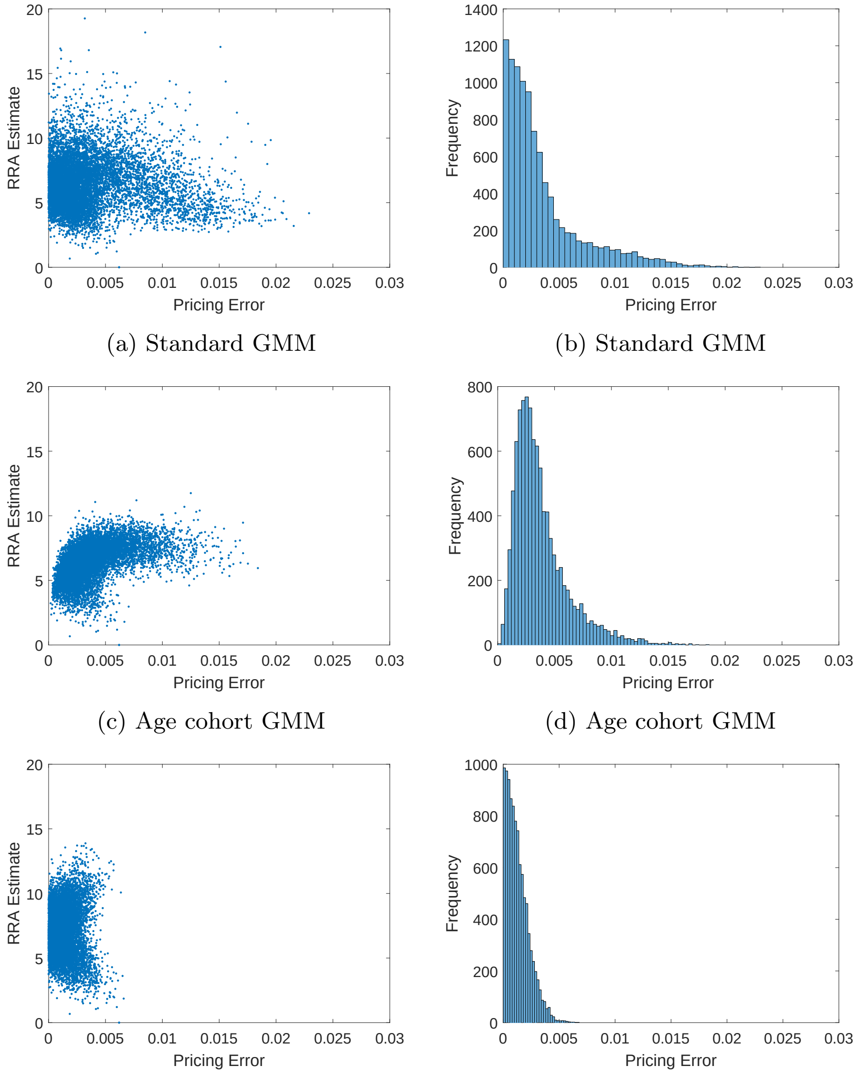


(f) Representative-agent GMM

Figure 4.1: GMM estimation of conditional model (assets: dividend claim; instruments: P/D ratio). Left: scatter plot of γ estimates and pricing errors. Right: histogram of pricing errors. $T = 500$.

the true model, while age cohort under-rejects and RA has the correct size. Overall, this exercise suggests that the nonexistent moments lead to low, overfit pricing errors with high γ estimates in many instances and excessively high pricing errors in others. On net, this leads standard GMM to over-reject both the true parameter and model.

Why is the γ estimate so imprecise with the standard GMM? Spurious troughs in the GMM criterion seem to be the cause. For the exactly identified, conditional, and unconditional specifications, in 1096, 1079, and 254 out of 10,000 simulations (respectively), the standard GMM criterion has multiple



(a) Standard GMM

(b) Standard GMM

(c) Age cohort GMM

(d) Age cohort GMM

(e) Representative-agent GMM

(f) Representative-agent GMM

Figure 4.2: GMM estimation of unconditional model (assets: dividend and consumption claims; instruments: none). Left: scatter plot of γ estimates and pricing errors. Right: histogram of pricing errors. $T = 500$.

inflection points, yielding one trough near the true γ and one or more in the moment nonexistence range. It seems nonexistent moments may introduce spurious troughs, and in some instances, a spurious one is closest to zero. Figure 4.3 illustrates this scenario. In this figure, there is a trough at $\gamma = 5.98$, which is close to the true value (7) but not the global minimum. The other trough at $\gamma = 18.1$ is the global minimum. In contrast, the age cohort GMM criterion has multiple troughs in 6, 6, and 0 out of 10,000 simulations.

Table 4.1: Finite sample properties

GMM	Standard			Age cohort			RA		
Sample size	100	300	500	100	300	500	100	300	500
Conditional model									
(assets: dividend claim only; instruments: price-dividend ratio)									
Bias	-1.39	-.69	-.23	-2.02	-1.05	-.67	.48	.048	.043
SE100	10.9	7.44	6.36	3.09	2.17	1.82	4.88	2.72	2.10
MAE	3.23	2.32	2.04	2.41	1.61	1.33	3.85	2.18	1.67
RMSE	3.96	2.94	2.69	3.13	2.11	1.72	4.95	2.75	2.11
Unconditional model									
(assets: both dividend and consumption claims; instruments: none)									
Bias	-1.79	-.92	-.56	-1.84	-.95	-.61	.37	.066	.064
SE100	6.42	4.13	3.28	2.64	1.87	1.56	4.06	2.33	1.80
MAE	2.73	1.91	1.57	2.15	1.42	1.16	3.28	1.88	1.45
RMSE	3.29	2.37	1.98	2.81	1.85	1.49	4.15	2.38	1.82

Note: SE100 denotes the mean of standard errors of $\widehat{\gamma}$ truncated at 100. MAE is the mean absolute error (average of $|\widehat{\gamma} - \gamma|$ across simulations). RMSE is the root mean squared error (square root of the average of $|\widehat{\gamma} - \gamma|^2$ across simulations).

Table 4.2: Type I error probabilities

GMM	Standard			Age cohort			RA		
Sample size	100	300	500	100	300	500	100	300	500
Conditional model									
(assets: dividend claim only; instruments: price-dividend ratio)									
Reject $\gamma = 7$	.025	.075	.076	.038	.040	.040	.025	.050	.050
Reject model	.13	.12	.16	.046	.004	.000	.086	.060	.059
Unconditional model									
(assets: both dividend and consumption claims; instruments: none)									
Reject $\gamma = 7$	.046	.092	.087	.047	.044	.042	.040	.052	.052
Reject model	.12	.13	.13	.036	.011	.011	.059	.050	.052

Note: $\gamma = 7$: t-test. Model: χ^2 test of overidentifying restrictions. Significance level: .05.

4.3 Type II error

To study the possibility of a Type II error, we use the model-simulated consumption data in conjunction with *false* asset returns data. More precisely, we generate a random permutation of the time index, and we use the model asset returns for this time index coupled with the consumption data of the calendar time. Because the equity premium is the same (2.11%) as with the true process and because the stochastic discount factor is always positive, the independence of the SDF and the excess stock returns (which holds by construction) implies that in large samples the moment condition *does not* hold.

From top to bottom, Figure 4.4 shows the conditional model results for standard GMM, age cohort GMM, and representative-agent GMM, respectively.

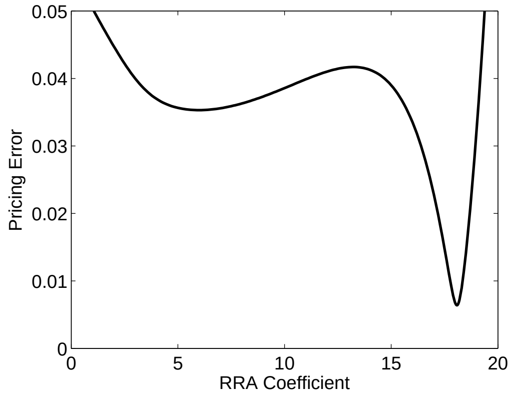


Figure 4.3: Normalized mean squared pricing error from a simulation with two troughs. Model: standard GMM. Specification: conditional (assets: dividend claim; instruments: P/D ratio).

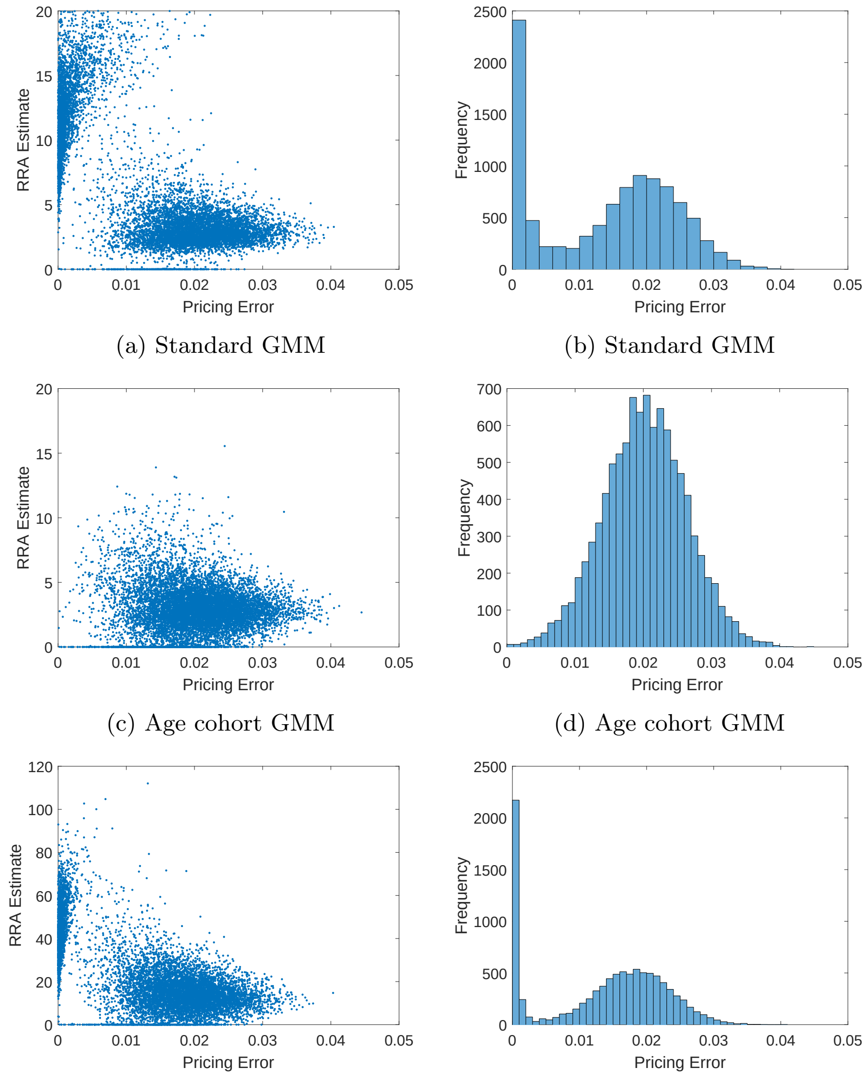
The left panels show the scatter plot of simulated γ estimates and normalized mean squared pricing errors for 10,000 simulations. The right panels show the histogram of the pricing errors.

Figure 4.4b shows the histogram of the pricing errors estimated by standard GMM. In 1877 out of 10,000 simulations, the pricing errors are within 10^{-3} of zero! With age cohort GMM (Figure 4.4d), in contrast, only 7 of 10,000 simulations yield pricing errors within 10^{-3} of zero. Also, the age cohort histogram is centered on about 2%, exactly as one would expect since the true equity premium is 2.11% and the pricing errors are normalized. Oddly, however, the standard GMM pricing error histogram is bimodal, with one peak at 2% and the other at zero. Moreover, as we see in Figures 4.4a and 4.4c, the spurious mode at zero is driven by upwardly biased estimates in the nonexistence range. This behavior is odd but perhaps unsurprising given the findings of Toda and Walsh (2015): the bootstrapped scatter plots and histograms of that analysis display *precisely the same pattern*!

Figure 4.5 is the same as 4.4 but with the unconditional specifications. As with Type I errors, switching from the conditional to unconditional model, which has two assets but no instruments, mitigates somewhat the spurious mass at zero for standard GMM (and for the RA model as well). However, Figures 4.5b and 4.5f still exhibit excess mass at zero, relative to age cohort, corresponding to high γ estimates.

Thus the standard GMM seems to lead to Type II errors (incorrect non-rejection of a false model) due to excessively low pricing errors. We can see formally the low power of standard GMM by comparing the histograms of the pricing errors of the true and false models. For example, under the null (consumption and return data generated from the true model), for standard GMM with the conditional specification ($T = 500$) the 95 percentile of the pricing error is .0139. Since the number of pricing errors larger than .0139 with the false model is 5756 out of 10,000 simulations, the rejection rate (power) is only 57.6%. On the other hand, for age cohort GMM, it is 91.6%.

Table 4.3 shows the Type II error probability (1 minus rejection rate or power) for each model using various tests. The row labeled "Pricing errors"



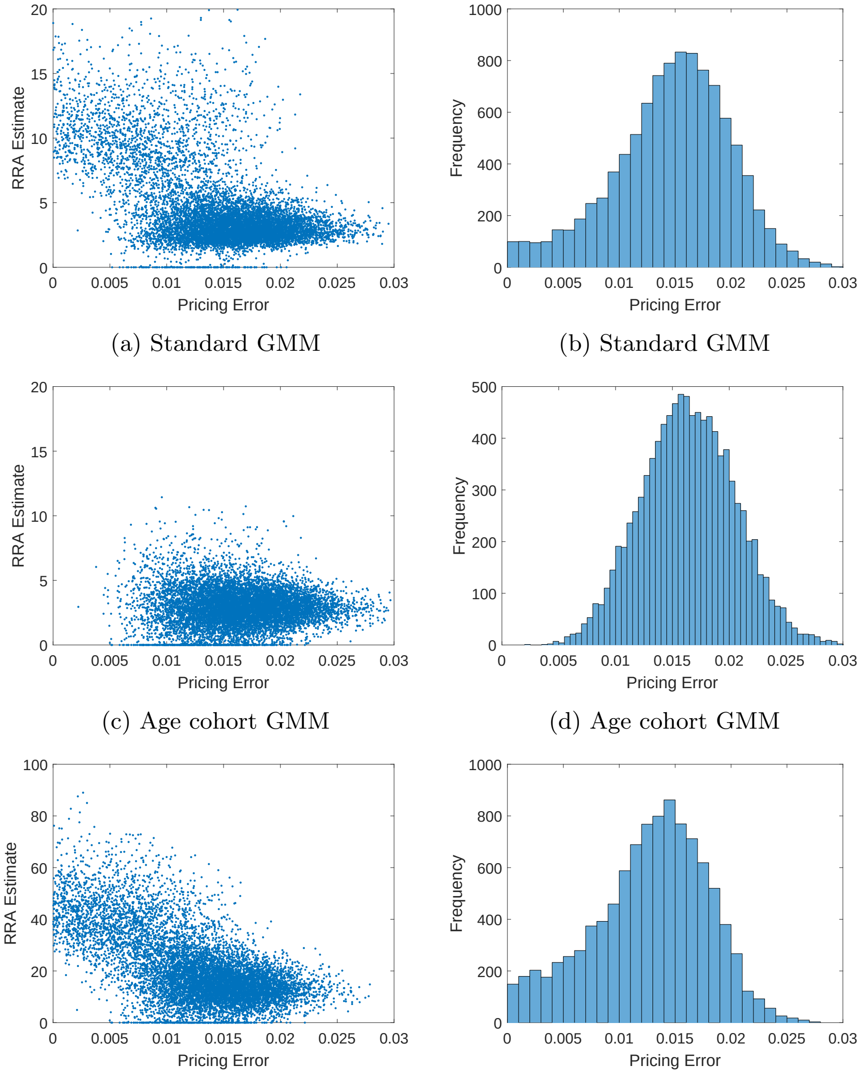
(a) Standard GMM (b) Standard GMM

(c) Age cohort GMM (d) Age cohort GMM

(e) Representative-agent GMM (f) Representative-agent GMM

Figure 4.4: GMM estimation of conditional model (assets: dividend claim; instruments: P/D ratio) with false stock returns. Left: scatter plot of γ estimates and pricing errors. Right: histogram of pricing errors. $T = 500$.

shows the result for the exact test just described using pricing errors. "Asymptotic χ^2 test" uses the χ^2 statistic from the first stage GMM and the critical value from the asymptotic distribution. "Exact χ^2 test" uses the same χ^2 statistic but obtains the critical value as the simulated 95 percentile under the null. The standard GMM is a disaster. Even with $T = 500$ and using the unconditional specification, the Type II error probability is 18 to 30 percent, depending on the test. With the conditional model, the range is 36 to 99 percent! The representative-agent GMM is similar when using the χ^2 statistic, although the performance is better when using the pricing errors probably because they are so



(a) Standard GMM (b) Standard GMM

(c) Age cohort GMM (d) Age cohort GMM

(e) Representative-agent GMM (f) Representative-agent GMM

Figure 4.5: GMM estimation of unconditional model (assets: dividend and consumption claims; instruments: none) with false stock returns. Left: scatter plot of γ estimates and pricing errors. Right: histogram of pricing errors. $T = 500$.

small under the correct null. In contrast, the age cohort GMM has much higher power with respect to the pricing error and exact χ^2 test: the Type II error probability is around 3 to 17 percent, depending on the specification and test. Age cohort GMM, however, performs poorly with respect to the asymptotic χ^2 test.

In summary, a plausible explanation for the emergence of the spurious peak at zero is that the fat tails mechanically aid in zeroing out the pricing errors. Indeed, using the sample versions of nonexistent moments seems to cause over-

Table 4.3: Type II error probabilities

GMM	Standard			Age cohort			RA		
Sample size	100	300	500	100	300	500	100	300	500
Conditional model									
(assets: dividend claim only; instruments: price-dividend ratio)									
Pricing errors	.79	.58	.42	.71	.28	.08	.35	.15	.10
Asymptotic χ^2 test	.75	.53	.36	.86	.87	.75	.76	.54	.37
Exact χ^2 test	.96	.98	.99	.85	.43	.17	.83	.57	.38
Unconditional model									
(assets: both dividend and consumption claims; instruments: none)									
Pricing errors	.74	.44	.24	.64	.18	.029	.33	.10	.057
Asymptotic χ^2 test	.68	.32	.18	.89	.78	.44	.73	.39	.24
Exact χ^2 test	.85	.55	.30	.86	.44	.14	.76	.39	.25

Note: The table shows, with respect to different tests, the probability of failing to reject that the model explains false, randomly generated returns. See text for explanations of the various tests. Significance level: .05.

fitting of models. This conjecture seems to hold for the representative-agent model as well. In this case, aggregate consumption growth is lognormal, so the tails are thin. However, by raising a lognormal variable to a high power, we can get tails that are quite fat. As in the case with standard GMM, the histogram of representative-agent GMM in Figure 4.4f shows a bimodal pattern. In 2172 out of 10,000 simulations, the pricing errors are within 10^{-3} of zero, and the spurious mode at zero is driven by upwardly biased estimates of γ around 20 to 80 according to the scatter plot in Figure 4.4e.

4.4 Source of bimodal pricing errors

What is the source of bimodality in the pricing error, with a spurious peak at zero?[18] We can provide an intuitive explanation as follows. Consider the GMM estimation of the representative-agent model with no instruments (single equation). Then the GMM estimator is the solution of

$$\frac{1}{T}\sum_{t=1}^{T}(C_t/C_{t-1})^{-\gamma}(R_t^s - R_{t-1}^f) = 0. \tag{4.1}$$

For notational simplicity, let $G_t = C_t/C_{t-1}$ be the aggregate consumption growth and $X_t = R_t^s - R_{t-1}^f$ be the excess stock market return. Furthermore, relabel time so that $G_1 \leq G_2 \leq \cdots \leq G_T$. Then (4.1) becomes

$$\frac{1}{T}\sum_{t=1}^{T} G_t^{-\gamma} X_t = 0. \tag{4.2}$$

[18]A number of previous studies have shown that instrumental variable estimation and fat tails may cause bimodality in test statistic distributions. See Nelson and Startz (1990) and Fiorio et al. (2010), for example. Andrews and Cheng (2012) show that weak identification can cause bimodal or skewed estimator and test statistic distributions.

Since $\{G_t\}_{t=1}^T$ is sorted in ascending order and $\gamma > 0$, we have $G_1^{-\gamma} \gg G_2^{-\gamma} \gg \cdots \gg G_T^{-\gamma}$. Hence the first two terms dominate the others, and (4.2) becomes

$$G_1^{-\gamma} X_1 + G_2^{-\gamma} X_2 = 0 \tag{4.3}$$

approximately. But provided that X_1, X_2 have opposite signs and $|X_2| > |X_1|$, (4.3) has a solution

$$\gamma = \frac{\log(-X_2/X_1)}{\log(G_2/G_1)} > 0. \tag{4.4}$$

Since G_t's are sorted in ascending order, G_1 and G_2 are relatively close to each other, so $\log(G_2/G_1)$ is a small positive number. Therefore the γ estimate in (4.4) will typically be a large number. Note that this argument holds regardless of whether the model is true or false. If asset returns are completely random, we would expect that we can make the pricing error close to zero with probability $\Pr(-X_2/X_1 > 1)$, which will be the probability of Type II errors.

The same argument holds for the standard GMM estimation with MU SDF. Recall that the MU SDF is defined by

$$\widehat{m}_t^{\mathrm{MU}}(\gamma) = \frac{\frac{1}{I}\sum_{i=1}^I c_{it}^{-\gamma}}{\frac{1}{I}\sum_{i=1}^I c_{i,t-1}^{-\gamma}}.$$

When cross-sectional consumption has fat tails, the terms corresponding to the minimum consumption at each period dominate, and we have

$$\widehat{m}_t^{\mathrm{MU}}(\gamma) \approx \frac{\frac{1}{I}(\min_i c_{it})^{-\gamma}}{\frac{1}{I}(\min_i c_{i,t-1})^{-\gamma}} = \left(\frac{\min_i c_{it}}{\min_i c_{i,t-1}}\right)^{-\gamma}.$$

Thus the same argument holds by replacing C_t/C_{t-1} in (4.1) by $\min_i c_{it}/\min_i c_{i,t-1}$. In particular, the MU γ estimates from standard GMM will be biased upwards as in Figures 4.1a and 4.4a because the γ given by (4.4) tends to be large. In Section 2.2 we showed formally that GMM minimizes a random function $Z(\gamma)^2$, and we now have an intuitive explanation for this phenomenon: the GMM estimate depends on random outlier draws, even when T is large.

Now we can see what the age cohort GMM achieves. For a false model, the pricing error is spuriously set to zero at the γ given by (4.4). Note that this γ depends on the value of G_2/G_1, the fraction between the two smallest observations. With the MU SDF, G_t corresponds to $\min_i c_{it}/\min_i c_{i,t-1}$. By dividing agents into age cohorts, the value of $\min_i c_{it}/\min_i c_{i,t-1}$ for each cohort will in general be distinct. Therefore except for by chance, it would not be possible to set the pricing errors simultaneously zero across age cohorts. Only if the model is true can we set the pricing errors simultaneously zero at the true γ. This gives age cohort GMM statistical power higher than that of standard GMM. A similar argument holds for the unconditional specification with two assets since the signs of $R_t^c - R_{t-1}^f$ and $R_t^d - R_{t-1}^f$ will often not be the same. Standard GMM, in contrast, may zero the pricing error at the arbitrary γ from (4.4) whether or not returns are generated from the true model.

5 Conclusion

In order to use GMM to estimate and test heterogeneous-agent consumption-based asset pricing models, many studies have employed the technique of averaging across the Euler equations of individual households. We simulate asset prices and a fat tailed consumption distribution from a tractable incomplete-market dynamic general equilibrium model and show in a Monte Carlo study that there are potential pitfalls to this practice of averaging: in the presence of fat tails in the cross-section, the resulting GMM criterion may contain sample analogs of nonexistent moments, which diverge in large samples. We establish that fat tails in consumption create over-rejection of true models/parameters and Type II errors (non-rejection of incorrect models) in the standard aggregated Euler equation GMM estimation of the relative risk aversion coefficient. The "age cohort" estimation method suggested in Toda and Walsh (2015) appears to mitigate these problems. Our broad message is that standard inference methods may be invalid in settings prone to power laws.

When should we worry about fat tails, and what should we do to avoid spurious estimation? Our Monte Carlo exercise sheds some light on these issues. First, even the representative-agent model (which does not have fat tails) is prone to spurious estimation by raising a positive random variable (here consumption growth) to a high power, which makes the tails fatter. So one should be careful when estimating a model that involves a power function. Second, spurious estimation seems to result from minimizing the sample GMM criterion by canceling the two outliers with opposite signs. Since the location of this spurious trough is random, estimating an overidentified model will likely mitigate the problem. Finally, when in doubt we can always conduct a bootstrap exercise, for example the stationary bootstrap of Politis and Romano (1994). According to the findings of Toda and Walsh (2015), a bimodal histogram of bootstrapped GMM criterions suggests spurious estimation.

A Asset prices

Since agents have identical homothetic preferences and all shocks are multiplicative (additive in logs), it is known that even if there are arbitrarily many assets, as long as the payoffs of the assets do not depend on idiosyncratic shocks, there will be no trade in assets in equilibrium, that is, the equilibrium is autarky (Constantinides and Duffie, 1996; Krueger and Lustig, 2010; Toda, 2014). Thus individual consumption c_{it} equals individual endowment y_{it}. By the first-order condition for the stock, we have

$$c_{it}^{-\gamma} P_t^d = \tilde{\beta}\,\mathrm{E}_t[c_{i,t+1}^{-\gamma}(P_{t+1}^d + D_{t+1})], \tag{A.1}$$

where P_t^d is the price of the dividend claim and $\tilde{\beta} = \beta(1-\delta)$ is the effective discount factor. Dividing both sides by $c_{it}^{-\gamma} D_t$ and defining the price-dividend ratio in state x_t by $V_d(x_t) := P_t^d / D_t$, we obtain

$$\begin{aligned} V_d(x_t) &= \tilde{\beta}\,\mathrm{E}_t[(c_{i,t+1}/c_{it})^{-\gamma}(D_{t+1}/D_t)(V_d(x_{t+1})+1)] \\ &= \tilde{\beta}\,\mathrm{E}_t\left[\exp\left(-\gamma\log(C_{t+1}/C_t) - \gamma\varepsilon_{i,t+1} + \log(D_{t+1}/D_t)\right)(V_d(x_{t+1})+1)\right]. \end{aligned}$$

Letting $v_d = (-\gamma, 1)'$ and using the fact that ε_{it} is i.i.d., we obtain

$$\begin{aligned} V_d(x_t) &= \tilde{\beta}\,\mathrm{E}_t[\exp(-\gamma\varepsilon_{i,t+1})]\,\mathrm{E}_t\left[\exp(v_d' x_{t+1})(V_d(x_{t+1})+1)\right] \\ &= \tilde{\beta}\mathrm{e}^{\frac{1}{2}\gamma(\gamma+1)\sigma^2}\,\mathrm{E}_t[\exp(v_d' x_{t+1})(V_d(x_{t+1})+1)], \end{aligned} \tag{A.2}$$

where we have used

$$\begin{aligned}\mathrm{E}[\exp(-\gamma\varepsilon)] &= \int_{-\infty}^{\infty} \mathrm{e}^{-\gamma x}\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x+\sigma^2/2)^2}{2\sigma^2}}\mathrm{d}x \\ &= \int_{-\infty}^{\infty} \mathrm{e}^{\frac{1}{2}\gamma(\gamma+1)\sigma^2}\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x+(\gamma+1/2)\sigma^2)^2}{2\sigma^2}}\mathrm{d}x = \mathrm{e}^{\frac{1}{2}\gamma(\gamma+1)\sigma^2}\end{aligned}$$

if $\varepsilon \sim N(-\sigma^2/2, \sigma^2)$. When $\{x_t\}$ follows a VAR(1) process (3.1), Burnside (1998) iterates (A.2) and obtains a closed-form solution as follows. Let

$$\begin{aligned}\tilde{\Sigma} &= (I-A)^{-1}\Sigma(I-A')^{-1}, \\ \Sigma_n &= \sum_{k=1}^{n} A^k\tilde{\Sigma}(A')^k, \\ B_n &= \sum_{k=1}^{n} A^k = A(I-A^n)(I-A)^{-1}, \\ \Omega_n &= n\tilde{\Sigma} - B_n\tilde{\Sigma} - \tilde{\Sigma}B_n' + \Sigma_n.\end{aligned}$$

Then we have

$$V_d(x) = \sum_{n=1}^{\infty}\tilde{\beta}^n \exp\left(\left(\frac{1}{2}\gamma(\gamma+1)\sigma^2 + v_d'g\right)n + v_d'B_n(x-g) + \frac{1}{2}v_d'\Omega_n v_d\right). \tag{A.3}$$

It is easy to show that this series converges if and only if

$$\tilde{\beta}\exp\left(\frac{1}{2}\gamma(\gamma+1)\sigma^2 + v_d'g + \frac{1}{2}v_d'\tilde{\Sigma}v_d\right) < 1. \tag{A.4}$$

Since $v_d = (-\gamma, 1)'$, inside of the exponential is a quadratic function in each of σ and γ. Therefore in order for an equilibrium to exist, the idiosyncratic volatility σ or risk aversion γ cannot be too high.

We can compute the asset returns as follows. Let $x_t = (x_{1t}, x_{2t})'$. Then the dividend growth is $D_{t+1}/D_t = \mathrm{e}^{x_{2,t+1}}$, and the stock return is

$$R_{t+1}^d = \frac{P_{t+1}^d + D_{t+1}}{P_t^d} = \frac{(P_{t+1}^d/D_{t+1}+1)(D_{t+1}/D_t)}{P_t^d/D_t} = \frac{V_d(x_{t+1})+1}{V_d(x_t)}\mathrm{e}^{x_{2,t+1}}. \tag{A.5}$$

We can compute the return to the consumption claim similarly by computing $V_c(x_t)$ as in (A.2) with $v_c = (1-\gamma, 0)'$ instead of v_d and using (A.5) to define R_{t+1}^c with $V_c, x_{1,t+1}$ instead of $V_d, x_{2,t+1}$. The calculation of the risk-free rate R_t^f is similar. Letting $v_f = (-\gamma, 0)'$, by the Euler equation we have

$$\begin{aligned}\frac{1}{R_t^f} &= \tilde{\beta}\,\mathrm{E}_t[(c_{i,t+1}/c_{it})^{-\gamma}] = \tilde{\beta}\,\mathrm{E}_t[\exp(-\gamma\log(C_{t+1}/C_t) - \gamma\varepsilon_{i,t+1})] \\ &= \tilde{\beta}\exp\left(\frac{1}{2}\gamma(\gamma+1)\sigma^2 + v_f'(g + A(x_t - g)) + \frac{1}{2}v_f'\Sigma v_f\right).\end{aligned} \tag{A.6}$$

References

Sule Alan, Orazio P. Attanasio, and Martin Browning. Estimating Euler equations with noisy data: Two exact GMM estimators. *Journal of Applied Econometrics*, 24(2):309–324, March 2009. doi:10.1002/jae.1037.

Joseph G. Altonji and Lewis M. Segal. Small-sample bias in GMM estimation of covariance structures. *Journal of Business and Economic Statistics*, 14(3): 353–366, July 1996. doi:10.1080/07350015.1996.10524661.

Donald W. K. Andrews and Xu Cheng. Estimation and inference with weak, semi-strong and strong identification. *Econometrica*, 80(5):2153–2211, September 2012. doi:10.3982/ECTA9456.

Pierluigi Balduzzi and Tong Yao. Testing heterogeneous-agent models: An alternative aggregation approach. *Journal of Monetary Economics*, 54(2): 369–412, March 2007. doi:10.1016/j.jmoneco.2005.08.021.

Parantap Basu, Andrei Semenov, and Kenji Wada. Uninsurable risk and financial market puzzles. *Journal of International Money and Finance*, 30(6): 1055–1089, October 2011. doi:10.1016/j.jimonfin.2011.05.012.

Erich Battistin, Richard Blundell, and Arthur Lewbel. Why is consumption more log normal than income? Gibrat's law revisited. *Journal of Political Economy*, 117(6):1140–1154, December 2009. doi:10.1086/648995.

Marie-Claude Beaulieu, Jean-Marie Dufour, and Lynda Khalaf. Asset-pricing anomalies and spanning: Multivariate and multifactor tests with heavy-tailed distributions. *Journal of Empirical Finance*, 17(4):763–782, September 2010. doi:10.1016/j.jempfin.2010.03.001.

Jess Benhabib, Alberto Bisin, and Shenghao Zhu. The distribution of wealth and fiscal policy in economies with finitely lived agents. *Econometrica*, 79(1): 123–157, January 2011. doi:10.3982/ECTA8416.

Jess Benhabib, Alberto Bisin, and Shenghao Zhu. The distribution of wealth in the Blanchard-Yaari model. *Macroeconomic Dynamics*, 20:466–481, March 2016. doi:10.1017/S1365100514000066.

István Berkes and Lajos Horváth. The rate of consistency of the quasi-maximum likelihood estimator. *Statistics and Probability Letters*, 61(2):133–143, January 2003. doi:10.1016/S0167-7152(02)00342-5.

Olivier J. Blanchard. Debt, deficits, and finite horizons. *Journal of Political Economy*, 93(2):223–247, April 1985.

Alon Brav, George M. Constantinides, and Christopher C. Geczy. Asset pricing with heterogeneous consumers and limited participation: Empirical evidence. *Journal of Political Economy*, 110(4):793–824, August 2002. doi:10.1086/340776.

Douglas T. Breeden. An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics*, 7(3):265–296, September 1979. doi:10.1016/0304-405X(79)90016-3.

Leo Breiman. *Probability*. Addison Wesley, Reading, MA, 1968.

Matthew Brzozowski, Martin Gervais, Paul Klein, and Michio Suzuki. Consumption, income, and wealth inequality in Canada. *Review of Economic Dynamics*, 13(1):52–75, January 2010. doi:10.1016/j.red.2009.10.006.

Craig Burnside. Solving asset pricing models with Gaussian shocks. *Journal of Economic Dynamics and Control*, 22(3):329–340, March 1998. doi:10.1016/S0165-1889(97)00075-4.

Craig Burnside. Identification and inference in linear stochastic discount factor models with excess returns. *Journal of Financial Econometrics*, 14(2):295–330, 2016. doi:10.1093/jjfinec/nbv018.

Christopher Dixon Carroll. Death to the log-linearized consumption Euler equation! (and very poor health to the second-order approximation). *Advances in Macroeconomics*, 1(1), 2001. doi:10.2202/1534-6013.1003.

Todd E. Clark. Small-sample properties of estimators of nonlinear models of covariance structure. *Journal of Business and Economic Statistics*, 14(3): 367–373, July 1996. doi:10.1080/07350015.1996.10524662.

Daren B. H. Cline. Convolution tails, product tails and domains of attraction. *Probability Theory and Related Fields*, 72(4):529–557, 1986. doi:10.1007/BF00344720.

John H. Cochrane. *Asset Pricing*. Princeton University Press, Princeton, NJ, second edition, 2005.

Timothy Cogley. Idiosyncratic risk and the equity premium: Evidence from the consumer expenditure survey. *Journal of Monetary Economics*, 49(2): 309–334, March 2002. doi:10.1016/S0304-3932(01)00106-4.

George M. Constantinides and Darrell Duffie. Asset pricing with heterogeneous consumers. *Journal of Political Economy*, 104(2):219–240, April 1996.

George M. Constantinides and Anisha Ghosh. Asset pricing with countercyclical household consumption risk. NBER Working Paper 20110, 2014. URL `http://www.nber.org/papers/w20110`.

Richard A. Davis and Tailen Hsing. Point process and partial sum convergence for weakly dependent random variables with infinite variance. *Annals of Probability*, 23(2):879–917, 1995. doi:10.1214/aop/1176988294.

Angus Deaton and Christina Paxson. Intertemporal choice and inequality. *Journal of Political Economy*, 102(3):437–467, June 1994.

Richard Durrett. *Probability: Theory and Examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, New York, fourth edition, 2010.

Paul Embrechts and Charles M. Goldie. Closure and factorization properties of subexponential and related distributions. *Journal of the Australian Mathematical Society (Series A)*, 29(02):243–256, March 1980. doi:10.1017/S1446788700021224.

Carlo V. Fiorio, Vassilis A. Hajivassiliou, and Peter C. B. Phillips. Bimodal t-ratios: The impact of thick tails on inference. *Econometrics Journal*, 13(2): 271–289, July 2010. doi:10.1111/j.1368-423X.2010.00315.x.

Xavier Gabaix. Power laws in economics and finance. *Annual Review of Economics*, 1:255–293, 2009. doi:10.1146/annurev.economics.050708.142940.

John Geweke. A note on some limitations of CRRA utility. *Economics Letters*, 71(3):341–345, June 2001. doi:10.1016/S0165-1765(01)00391-3.

Nikolay Gospodinov, Raymond Kan, and Cesare Robotti. Misspecification-robust inference in linear asset-pricing models with irrelevant risk factors. *Review of Financial Studies*, 27(7):2139–2170, 2014. doi:10.1093/rfs/hht135.

Sanford J. Grossman and Robert J. Shiller. The determinants of the variability of stock market prices. *American Economic Review Papers and Proceedings*, 71(2):222–227, May 1981.

Sanford J. Grossman, Angelo Melino, and Robert J. Shiller. Estimating the continuous-time consumption-based asset-pricing model. *Journal of Business and Economic Statistics*, 5(3):315–327, 1987. doi:10.1080/07350015.1987.10509594.

Fatih Guvenen. A parsimonious macroeconomic model for asset pricing. *Econometrica*, 77(6):1711–1750, November 2009. doi:10.3982/ECTA6658.

Peter Hall and Qiwei Yao. Inference in ARCH and GARCH models with heavy-tailed errors. *Econometrica*, 71(1):285–317, January 2003. doi:10.1111/1468-0262.00396.

Lars Peter Hansen and Kenneth J. Singleton. Stochastic consumption, risk aversion, and the temporal behavior of asset returns. *Journal of Political Economy*, 91(2):249–265, April 1983.

Lars Peter Hansen, John Heaton, and Amir Yaron. Finite-sample properties of some alternative GMM estimators. *Journal of Business and Economic Statistics*, 14(3):262–280, July 1996. doi:10.1080/07350015.1996.10524656.

John Heaton and Deborah J. Lucas. Evaluating the effects of incomplete markets on risk sharing and asset pricing. *Journal of Political Economy*, 104(3):443–487, June 1996.

Jonathan B. Hill. Robust estimation and inference for heavy tailed GARCH. *Bernoulli*, 21(3):1629–1669, 2015. doi:10.3150/14-BEJ616.

Jonathan B. Hill and Artem Prokhorov. GEL estimation for GARCH models with robust empirical likelihood inference. *Journal of Econometrics*, 190(1): 18–45, January 2016. doi:10.1016/j.jeconom.2015.09.001.

Raymond Kan and Chu Zhang. Two-pass tests of asset pricing models with useless factors. *Journal of Finance*, 54(1):203–235, February 1999a. doi:10.1111/0022-1082.00102.

Raymond Kan and Chu Zhang. GMM tests of stochastic discount factor models with useless factors. *Journal of Financial Economics*, 54(1):103–127, October 1999b. doi:10.1016/S0304-405X(99)00033-1.

Raymond Kan, Cesare Robotti, and Jay Shanken. Pricing model performance and the two-pass cross-sectional regression methodology. *Journal of Finance*, 68(6):2617–2649, December 2013. doi:10.1111/jofi.12035.

Yuichi Kitamura. Empirical likelihood methods in econometrics: Theory and practice. In Richard Blundell, Whitney Newey, and Torsten Persson, editors, *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress*, volume 3 of *Econometric Society Monographs*, chapter 7, pages 174–237. Cambridge University Press, New York, 2007.

Yuichi Kitamura and Michael Stutzer. An information-theoretic alternative to generalized method of moments estimation. *Econometrica*, 65(4):861–874, July 1997. doi:10.2307/2171942.

Narayana R. Kocherlakota. On tests of representative consumer asset pricing models. *Journal of Monetary Economics*, 26(2):285–304, October 1990. doi:10.1016/0304-3932(90)90024-X.

Narayana R. Kocherlakota. Testing the consumption CAPM with heavy-tailed pricing errors. *Macroeconomic Dynamics*, 1(3):551–567, September 1997. doi:10.1017/S136510059700401X.

Narayana R. Kocherlakota and Luigi Pistaferri. Asset pricing implications of Pareto optimality with private information. *Journal of Political Economy*, 117(3):555–590, June 2009. doi:10.1086/599761.

Samuel Kotz, Tomasz J. Kozubowski, and Krzysztof Podgórski. *The Laplace Distribution and Generalizations*. Birkhäuser, Boston, 2001.

Tom Krebs and Bonnie Wilson. Asset returns in an endogenous growth model with incomplete markets. *Journal of Economic Dynamics and Control*, 28(4): 817–839, January 2004. doi:10.1016/S0165-1889(03)00062-9.

Dirk Krueger and Hanno Lustig. When is market incompleteness irrelevant for the price of aggregate risk (and when is it not)? *Journal of Economic Theory*, 145(1):1–41, January 2010. doi:10.1016/j.jet.2009.10.005.

Dirk Krueger, Hanno Lustig, and Fabrizio Perri. Evaluating asset pricing models with limited commitment using household consumption data. *Journal of the European Economic Association*, 6(2-3):715–726, April-May 2008. doi:10.1162/JEEA.2008.6.2-3.715.

Per Krusell and Anthony A. Smith, Jr. Income and wealth heterogeneity in the macroeconomy. *Journal of Political Economy*, 106(5):867–896, October 1998. doi:10.1086/250034.

Robert E. Lucas, Jr. Asset prices in an exchange economy. *Econometrica*, 46 (6):1429–1445, November 1978. doi:10.2307/1913837.

Sydney C. Ludvigson. Advances in consumption-based asset pricing: Empirical tests. In George M. Constantinides, Milton Harris, and René M. Stultz, editors, *Handbook of the Economics of Finance*, volume 2, chapter 12, pages 799–906. Elsevier, Amsterdam, 2013. doi:10.1016/B978-0-44-459406-8.00012-3.

N. Gregory Mankiw. The equity premium and the concentration of aggregate shocks. *Journal of Financial Economics*, 17(1):211–219, September 1986. doi:10.1016/0304-405X(86)90012-7.

Rajnish Mehra and Edward C. Prescott. The equity premium: A puzzle. *Journal of Monetary Economics*, 15(2):145–161, March 1985. doi:10.1016/0304-3932(85)90061-3.

Charles R. Nelson and Richard Startz. Some further results on the exact small sample properties of the instrumental variable estimator. *Econometrica*, 58 (4):967–976, July 1990. doi:10.2307/2938359.

Whitney K. Newey and Richard J. Smith. Higher order properties of GMM and generalized empirical likelihood estimators. *Econometrica*, 72(1):219–255, January 2004. doi:10.1111/j.1468-0262.2004.00482.x.

Dimitris N. Politis and Joseph P. Romano. The stationary bootstrap. *Journal of the American Statistical Association*, 89(428):1303–1313, 1994. doi:10.1080/01621459.1994.10476870.

William J. Reed. The Pareto, Zipf and other power laws. *Economics Letters*, 74(1):15–19, December 2001. doi:10.1016/S0165-1765(01)00524-9.

William J. Reed. The Pareto law of incomes—an explanation and an extension. *Physica A*, 319(1):469–486, March 2003. doi:10.1016/S0378-4371(02)01507-8.

William J. Reed and Murray Jorgensen. The double Pareto-lognormal distribution—a new parametric model for size distribution. *Communications in Statistics—Theory and Methods*, 33(8):1733–1753, 2004. doi:10.1081/STA-120037438.

Sidney I. Resnick. *Extreme Values, Regular Variation, and Point Processes*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2008.

Makoto Saito. A simple model of incomplete insurance: The case of permanent shocks. *Journal of Economic Dynamics and Control*, 22(5):763–777, May 1998. doi:10.1016/S0165-1889(97)00077-8.

Alexi Savov. Asset pricing with garbage. *Journal of Finance*, 66(1):177–201, February 2011. doi:10.1111/j.1540-6261.2010.01629.x.

Lawrence D. W. Schmidt. Climbing and falling off the ladder: Asset pricing implications of labor market event risk. 2015. URL `http://ssrn.com/abstract=2471342`.

Andrei Semenov. Background risk in consumption and the equity risk premium. *Review of Quantitative Finance and Accounting*, 2016. doi:10.1007/s11156-016-0556-2.

Kjetil Storesletten, Chris I. Telmer, and Amir Yaron. Asset pricing with idiosyncratic risk and overlapping generations. *Review of Economic Dynamics*, 10(4):519–548, October 2007. doi:10.1016/j.red.2007.02.004.

George Tauchen. Statistical properties of generalized method-of-moments estimators of structural parameters obtained from financial market data. *Journal of Business and Economic Statistics*, 4(4):397–416, October 1986. doi:10.1080/07350015.1986.10509537.

Alexis Akira Toda. Incomplete market dynamics and cross-sectional distributions. *Journal of Economic Theory*, 154:310–348, November 2014. doi:10.1016/j.jet.2014.09.015.

Alexis Akira Toda. Asset prices and efficiency in a Krebs economy. *Review of Economic Dynamics*, 18(4):957–978, October 2015. doi:10.1016/j.red.2014.11.003.

Alexis Akira Toda. A note on the size distribution of consumption: More double Pareto than lognormal. *Macroeconomic Dynamics*, 2016. doi:10.1017/S1365100515000942.

Alexis Akira Toda and Kieran Walsh. The double power law in consumption and implications for testing Euler equations. *Journal of Political Economy*, 123(5):1177–1200, October 2015. doi:10.1086/682729.

Annette Vissing-Jørgensen. Limited asset market participation and the elasticity of intertemporal substitution. *Journal of Political Economy*, 110(4):825–853, August 2002. doi:10.1086/340782.