

# UC San Diego

## UC San Diego Previously Published Works

### Title

The Double Power Law in Consumption and Implications for Testing Euler Equations

### Permalink

<https://escholarship.org/uc/item/1jh2795s>

### Journal

Journal of Political Economy, 123(5)

### ISSN

0022-3808

### Authors

Toda, Alexis Akira

Walsh, Kieran

### Publication Date

2015-10-01

### DOI

10.1086/682729

Peer reviewed

# The Double Power Law in Consumption and Implications for Testing Euler Equations\*

Alexis Akira Toda<sup>†</sup>

*Department of Economics, University of California San Diego*

Kieran Walsh<sup>‡</sup>

*Darden School of Business, University of Virginia*

June 7, 2014

## Abstract

We provide evidence suggesting that the cross-sectional distributions of U.S. consumption and consumption growth obey the power law in both the upper and lower tails, with exponents approximately equal to 4. Consequently, high order moments are unlikely to exist, and the GMM estimation of Euler equations that employ cross-sectional moments may be inconsistent. Through bootstrap studies, we find that the power law appears to generate spurious non-rejection of heterogeneous-agent asset pricing models in explaining the equity premium. Dividing households into age groups, we propose an estimation approach which appears less susceptible to fat tail issues.

## 1 Introduction

There are many studies that use household level consumption data and historical financial asset returns to test the Euler equations of heterogeneous-agent models. Because micro consumption data contain measurement error and households participate in surveys for only short periods of time, this literature typically “aggregates” the Euler equations before estimating and testing. Consider the following example. Assume that households have identical additive constant relative risk aversion (CRRA) preferences  $E_0 \sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\gamma}}{1-\gamma}$ , where  $0 < \beta < 1$  is the discount factor and  $\gamma > 0$  is the relative risk aversion coefficient. Assuming interior solutions, the Euler equation

$$c_{it}^{-\gamma} = E [\beta c_{i,t+1}^{-\gamma} R_{t+1} \mid \mathcal{F}_{it}] \quad (1.1)$$

---

\*We benefited from comments by Donald Andrews, Tim Armstrong, Brendan Beare, Xiaohong Chen, David Childers, John Geanakoplos, Tony Smith, and seminar participants at Australian School of Business, UCSD, Yale, and the 17th International Conference on Macroeconomic Analysis and International Finance at the University of Crete. We thank two anonymous referees and the editor (Monika Piazzesi) for comments and suggestions that significantly improved the paper. AAT acknowledges the financial support from the Cowles Foundation, the Nakajima Foundation, and Yale University. KW acknowledges financial support from the Cowles Foundation and Yale University.

<sup>†</sup>Email: atoda@ucsd.edu

<sup>‡</sup>Email: walshk@darden.virginia.edu

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  $t$ . Let  $\mathcal{F}_t$  be the information set that contains only aggregate variables—in this example asset returns—and let  $\mathbf{E}_t$  denote the expectation conditional on  $\mathcal{F}_t$ . Taking the cross-sectional expectation, applying the law of iterated expectations on the Euler equation (1.1), and assuming that the cross-sectional moment  $\mathbf{E}_t[c_{it}^{-\gamma}]$  is finite, we obtain the moment condition

$$\mathbf{E}_t \left[ \mathbf{E}_t[c_{it}^{-\gamma}] - \beta \mathbf{E}_{t+1}[c_{i,t+1}^{-\gamma}] R_{t+1} \right] = 0.$$

Using the sample analogs, we can form the criterion

$$\left( \frac{1}{T} \sum_{t=1}^T \left( \frac{1}{I} \sum_{i=1}^I c_{it}^{-\gamma} - \beta R_{t+1} \frac{1}{I} \sum_{i=1}^I c_{i,t+1}^{-\gamma} \right) \right)^2$$

and minimize it to estimate the parameters  $(\beta, \gamma)$  by the generalized method of moments (GMM). This estimation method is consistent *if the cross-sectional moment  $\mathbf{E}_t[c_{it}^{-\gamma}]$  is finite*. But do the cross-sectional moments exist? If not, how should we estimate and test the model? If  $\mathbf{E}_t[c_{it}^{-\gamma}]$  does not exist, will the data reject the model, or might the moment condition be prone to over-fitting? These are the questions that we address in this paper.

This paper has three contributions. First, we document that the cross-sectional distributions of U.S. consumption and consumption growth data exhibit fat tails. More precisely, consumption and consumption growth seem to obey the power law in both the upper and lower tails with exponents approximately equal to 4. If these power laws hold, the cross-sectional moments of consumption  $\mathbf{E}_t[c_{it}^\eta]$  and consumption growth  $\mathbf{E}_t[(c_{it}/c_{i,t-1})^\eta]$  do not exist when  $|\eta|$  is large (that is, when  $\eta$  is well above or below zero), and the GMM estimation of the aggregated Euler equation is inconsistent.

Second, using a heterogeneous-agent consumption based asset pricing model as a laboratory and performing robustness checks such as dropping outliers and studying bootstrap samples, we find that the fat tails may cause spurious non-rejection of models.

Third, with the aim of mitigating the fat tail problem, we propose an alternative method for estimating and testing Euler equations. The intuition for our approach is the following. Because the Euler equation aggregation works for any conditioning variable, if we can find a variable such that the conditional consumption distribution does not have fat tails, then we can perform consistent GMM. In particular, we exploit the fact that the consumption distribution is approximately lognormal within age cohorts (Battistin et al., 2009) (within age group lognormality is also an implication of our theoretical model). Dividing the households into age cohorts, we form moment conditions corresponding to each age group and estimate and test an overidentified model. We find that this “age cohort GMM” appears to mitigate spurious non-rejection.

## 2 Literature and why fat tails matter for estimating and testing Euler equations

### 2.1 Euler equation aggregations

Consider the Euler equation

$$c_{it}^{-\gamma} = \mathbb{E} [\beta c_{i,t+1}^{-\gamma} R_{t+1} \mid \mathcal{F}_{it}], \quad (2.1)$$

which is the same as (1.1). 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). Dividing (2.1) by  $c_{it}^{-\gamma}$ , conditioning on aggregate variables  $\mathcal{F}_t$ , and applying the law of iterated expectations, we obtain

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

Thus ignoring the discount factor  $\beta$ , the  $-\gamma$ -th cross-sectional moment of consumption growth between time  $t$  and  $t + 1$ ,

$$m_{t+1}^{\text{IMRS}} = \mathbb{E}_{t+1} [(c_{i,t+1}/c_{it})^{-\gamma}],$$

is a valid stochastic discount factor (SDF), where IMRS stands for “intertemporal marginal rate of substitution”. For estimation, we can use the sample analog

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

The second approach is to average the Euler equation (2.1) directly, as in Balduzzi and Yao (2007). Taking the expectation of (2.1) with respect to  $\mathcal{F}_t$  and applying the law of iterated expectations, we obtain

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

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

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

Therefore ignoring  $\beta$ ,

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

is also a valid stochastic discount factor, where MU stands for “marginal utility”. Balduzzi and Yao (2007) argue that the MU SDF is less susceptible to measurement error. For estimation, we can use the sample analog

$$\widehat{m}_{t+1}^{\text{MU}} = \frac{\frac{1}{I} \sum_{i=1}^I c_{i,t+1}^{-\gamma}}{\frac{1}{I} \sum_{i=1}^I c_{it}^{-\gamma}}.$$

Kocherlakota and Pistaferri (2009) take a somewhat different approach. Instead of the Euler equation (2.1), they start from the inverse Euler equation, which holds in a private information setting when agents use insurance companies to achieve constrained Pareto optimal allocation. By a similar argument as in deriving the MU SDF, they obtain

$$m_{t+1}^{\text{PIPO}} = \frac{E_t[c_{it}^{\gamma}]}{E_{t+1}[c_{i,t+1}^{\gamma}]}$$

and use the sample analog

$$\widehat{m}_{t+1}^{\text{PIPO}} = \frac{\frac{1}{I} \sum_{i=1}^I c_{it}^{\gamma}}{\frac{1}{I} \sum_{i=1}^I c_{i,t+1}^{\gamma}}$$

for estimation. PIPO stands for “private information with Pareto optimality”.

As we will see in Section 2.3, the validity of the IMRS, MU, and PIPO stochastic discount factors relies on the existence of the cross-sectional moments  $E_t[(c_{it}/c_{i,t-1})^{-\gamma}]$ ,  $E_t[c_{it}^{-\gamma}]$ , and  $E_t[c_{it}^{\gamma}]$ , respectively. However, none of the above studies explicitly discusses the presence or implications of fat tails in the cross-sectional distribution of consumption or consumption growth.<sup>1</sup>

## 2.2 Empirical results

All of the above papers use household-level consumption data (Consumption Expenditure Survey, CEX) to empirically analyze and test the heterogeneous-agent, incomplete market approach.

Brav et al. (2002) and Cogley (2002) employ linearized versions of the sample analog of the IMRS stochastic discount factor. While the representative agent approach (Hansen and Singleton, 1983) considers only aggregate consumption (the cross-sectional mean of the consumption distribution), these papers try the cross-sectional mean, variance, and skewness of the consumption growth distribution. Brav et al. (2002) find that the IMRS SDF explains the equity premium for  $\gamma \approx 3.5$ , but Cogley (2002) finds that the equity premium is not explained for  $\gamma < 15$ . Vissing-Jørgensen (2002) follows a similar approach, but her focus is the estimation of the elasticity of intertemporal substitution and not the equity premium.

Balduzzi and Yao (2007) replicate the result of Brav et al. (2002) at the quarterly frequency but show that the IMRS SDF fails for monthly consumption growth. The main point of Balduzzi and Yao (2007) is that the MU SDF zeroes the pricing error (sample average of the moment condition errors) at

---

<sup>1</sup>Kocherlakota (1997) discusses the possibility of fat tails in aggregate consumption growth in the context of a representative agent model.

$\gamma \approx 10$  when they include only households with at least \$2,000 of financial assets. Also, assuming the consumption distribution is lognormal, they show that the MU SDF is a closed-form function of the change in mean and variance of the consumption distribution. This “BY” SDF performs similarly to MU.

While the above papers use CEX data from the early 1980s through the mid 1990s, Kocherlakota and Pistaferri (2009) analyze the longer sample from 1980 to 2004 and incorporate data from U.K. and Italy to perform overidentifying tests. In this longer sample, they reject the MU SDF even when restricting analysis to households that meet various asset thresholds. Their main result is that the PIPO SDF zeroes the pricing error at  $\gamma \approx 5$ . Also, imposing a common relative risk aversion  $\gamma$  across U.S., U.K., and Italy, overidentifying tests reject the representative agent (RA) and MU stochastic discount factors but not PIPO.

Table 1 below summarizes the literature of testing the heterogeneous-agent, incomplete market models. In summary, the literature had (i) generated mixed support for IMRS, (ii) confirmed MU (with less data) and then rejected it (with more recent data), and (iii) provided positive evidence for PIPO. See also Miller (1999) for a review of an earlier literature on the estimation of Euler equations with micro consumption data and Ludvigson (2013) for a more recent survey.

**Table 1.** Estimation of relative risk aversion  $\gamma$  and tests of stochastic discount factors in the literature.  $\checkmark$  (X) indicates support for (rejection of) an SDF. The number next to  $\checkmark$  is the estimate of  $\gamma$  when not rejected.  $Q$  ( $M$ ) indicates quarterly (monthly) consumption growth.

| Paper                              | Sample    | IMRS                            | MU              | PIPO           |
|------------------------------------|-----------|---------------------------------|-----------------|----------------|
| Brav et al. (2002)                 | 1982–1996 | $\checkmark$ 3.5                |                 |                |
| Cogley (2002)                      | 1980–1994 | X                               |                 |                |
| Balduzzi and Yao (2007)            | 1982–1995 | $Q$ : $\checkmark$ 5<br>$M$ : X | $\checkmark$ 10 |                |
| Kocherlakota and Pistaferri (2009) | 1980–2004 |                                 | X               | $\checkmark$ 5 |

### 2.3 Why fat tails matter

Why might fat tails in the consumption distribution create problems for GMM estimation?

First, consider the IMRS SDF. The relevant moment condition is

$$E [E_t[(c_{it}/c_{i,t-1})^{-\gamma}](R_t^s - R_t^b)] = 0,$$

where  $R_t^s$  is the stock return, and  $R_t^b$  is the bond return (so  $R_t^s - R_t^b$  is the excess return on stocks). GMM estimation of  $\gamma$  proceeds by forming the criterion

$$J_{T,\{I_t\}}^{\text{IMRS}}(\gamma) = \left( \frac{1}{T} \sum_{t=1}^T \frac{1}{I_t} \sum_{i=1}^{I_t} (c_{it}/c_{i,t-1})^{-\gamma} (R_t^s - R_t^b) \right)^2 \quad (2.2)$$

and minimizing it, where  $I_t$  is the number of households observed at time  $t$ . Assume that the cross-sectional moment  $E_t[(c_{it}/c_{i,t-1})^{-\gamma}]$  is finite only for  $\gamma \in [\underline{\gamma}, \bar{\gamma}]$ . Since for  $\gamma \notin [\underline{\gamma}, \bar{\gamma}]$  the sample moment  $\frac{1}{I_t} \sum_{i=1}^{I_t} (c_{it}/c_{i,t-1})^{-\gamma}$  tends to infinity as the number of households  $I_t$  tends to infinity, the GMM criterion (2.2) admits a large sample limit only if  $\gamma \in [\underline{\gamma}, \bar{\gamma}]$  and diverges to infinity otherwise.

Therefore unless the true value  $\gamma_0$  is in this range, we cannot estimate it by GMM.

Next consider the MU SDF. The relevant moment condition is

$$\mathbb{E} \left[ \frac{\mathbb{E}_t[c_{it}^{-\gamma}]}{\mathbb{E}_{t-1}[c_{i,t-1}^{-\gamma}]} (R_t^s - R_t^b) \right] = 0$$

and the GMM criterion is

$$J_{T,\{I_t\}}^{\text{MU}}(\gamma) = \left( \frac{1}{T} \sum_{t=1}^T \frac{\frac{1}{I_t} \sum_{i=1}^{I_t} c_{it}^{-\gamma}}{\frac{1}{I_{t-1}} \sum_{i=1}^{I_{t-1}} c_{i,t-1}^{-\gamma}} (R_t^s - R_t^b) \right)^2. \quad (2.3)$$

Assume that the cross-sectional moment  $\mathbb{E}_t[c_{it}^{-\gamma}]$  is finite only for  $\gamma \in [\underline{\gamma}, \bar{\gamma}]$ . Since for  $\gamma \notin [\underline{\gamma}, \bar{\gamma}]$  we have  $\frac{1}{I_t} \sum_{i=1}^{I_t} c_{it}^{-\gamma} / \frac{1}{I_{t-1}} \sum_{i=1}^{I_{t-1}} c_{i,t-1}^{-\gamma} \rightarrow \infty/\infty$  as the number of households  $I_{t-1}$  and  $I_t$  tend to infinity, the large sample limit of the GMM criterion (2.3) may not be well defined. One may still hope to find numbers  $N_{t-1}$  and  $N_t$  such that  $N_{t-1}/N_t = I_{t-1}/I_t$  and  $\frac{1}{N_t} \sum_{i=1}^{I_t} c_{it}^{-\gamma}$  converges to a finite value, yielding a well defined GMM criterion. However, the following theorem provides a negative answer.

**Theorem 2.1** (Feller (1946)). *Let  $X_1, X_2, \dots$  be i.i.d. with  $\mathbb{E}[|X_1|] = \infty$  and let  $S_n = X_1 + \dots + X_n$ . Let  $a_n$  be a sequence of positive numbers with  $a_n/n$  weakly increasing. Then  $\limsup_{n \rightarrow \infty} |S_n|/a_n = 0$  or  $\infty$  almost surely according as  $\sum_{n=1}^{\infty} P(|X_1| \geq a_n) < \infty$  or  $= \infty$ .*

Interpreting  $c_{it}^{-\gamma}$  as  $X_n$  and  $N_t$  as  $a_n$  in this theorem, it follows that the large sample limit of the GMM criterion will contain terms such as  $0/0$ ,  $0/\infty$ ,  $\infty/0$ , and  $\infty/\infty$ . Therefore if a finite limit exists, it must be zero. This argument shows that even if the true value  $\gamma_0$  belongs to the moment existence range  $[\underline{\gamma}, \bar{\gamma}]$ , it is not identified because values in the nonexistence range  $\gamma \notin [\underline{\gamma}, \bar{\gamma}]$  may set the GMM criterion to zero in the large sample limit. The same argument applies to PIPO.

In summary, if the cross-sectional moments of consumption or consumption growth do not exist, the large sample limit of the GMM criterion may not be well defined. Even if it is well defined it may be zero for  $\gamma \notin [\underline{\gamma}, \bar{\gamma}]$  distinct from the true value  $\gamma_0$ . Therefore standard GMM estimation is in general inconsistent unless (i) the true value  $\gamma_0$  belongs to the moment existence range  $[\underline{\gamma}, \bar{\gamma}]$ , and (ii) when estimating  $\gamma$  we restrict the search to the moment existence range  $[\underline{\gamma}, \bar{\gamma}]$ . This situation is quite problematic because the true value may not belong to the moment existence range, and even if it does, a priori we do not know the moment existence range. Additionally, GMM in this context may be prone to type II errors in which the model is incorrect but  $\gamma \notin [\underline{\gamma}, \bar{\gamma}]$  sets the criterion to zero.

### 3 Double power law in consumption

In this section we introduce the notion of the double power law and show both theoretically and empirically that the cross-sectional distribution of consumption exhibits fat tails.

### 3.1 Definition

A nonnegative random variable  $X$  obeys the *power law* (in the upper tail) with exponent  $\alpha > 0$  if

$$\lim_{x \rightarrow \infty} x^\alpha P(X > x) > 0$$

exists (Pareto, 1896; Mandelbrot, 1960; Gabaix, 2009). Recently, many economic variables have been shown to obey the power law also in the lower tail, meaning that

$$\lim_{x \rightarrow 0} x^{-\beta} P(X < x) > 0$$

exists for some exponent  $\beta > 0$ . Such phenomena have been found in city size (Giesen et al., 2010) and income (Toda, 2012). In this paper we say that  $X$  obeys the *double power law* if the power law holds in both the upper and the lower tails. If  $X$  obeys the double power law with exponents  $(\alpha, \beta)$ , then  $X^\eta$  obeys the double power law with exponents  $(\alpha/\eta, \beta/\eta)$  if  $\eta > 0$  and  $(-\beta/\eta, -\alpha/\eta)$  if  $\eta < 0$ . For example if  $\eta > 0$  we have

$$P(X^\eta > x) = P(X > x^{\frac{1}{\eta}}) \sim x^{-\frac{\alpha}{\eta}}$$

as  $x \rightarrow \infty$ , and other cases are similar. In this case the  $\eta$ -th moment  $E[X^\eta]$  exists if and only if  $-\beta < \eta < \alpha$ .

### 3.2 Theory

Why might the cross-sectional consumption distribution obey the double power law? To explore this possibility, consider an infinite horizon, continuous-time economy populated by a continuum of agents indexed by  $i \in I = [0, 1]$ . Each agent has the same additive CRRA preference

$$E_0 \int_0^\infty e^{-\rho t} \frac{c_t^{1-\gamma}}{1-\gamma} dt, \tag{3.1}$$

where  $\rho > 0$  is the time discount rate and  $\gamma > 0$  is the coefficient of relative risk aversion. Think of agents as entrepreneurs or dynasties operating private investment projects (*AK* technologies). Assume that capital invested in agent  $i$ 's project is subject to uninsurable idiosyncratic risk and evolves according to the geometric Brownian motion

$$dk_t/k_t = \mu dt + \sigma dB_{it},$$

where  $k_t$  is capital,  $\mu$  is the expected growth rate,  $\sigma > 0$  is the volatility, and  $B_{it}$  is a standard Brownian motion that is i.i.d. across agents. Assuming that agents can borrow or lend among each other using risk-free assets in zero net supply with an equilibrium risk-free rate  $r$ , the budget constraint of a typical agent becomes

$$dw_t = ((\mu\theta_t + r(1 - \theta_t))w_t - c_t)dt + \sigma\theta_t w_t dB_t, \tag{3.2}$$

where  $w_t$  is wealth,  $c_t$  is consumption, and  $\theta_t \geq 0$  is the fraction of wealth invested in the technology. The individual decision problem is to maximize utility (3.1) subject to the budget constraint (3.2). This problem is a classic Merton



(1971)-type optimal consumption-portfolio problem and therefore has a closed-form solution. Letting  $\varepsilon = 1/\gamma$  be the elasticity of intertemporal substitution, the solution is

$$\frac{c_t}{w_t} = m := \rho\varepsilon + (1 - \varepsilon) \left( r + \frac{(\mu - r)^2}{2\gamma\sigma^2} \right), \quad (3.3a)$$

$$\theta_t = \frac{\mu - r}{\gamma\sigma^2}. \quad (3.3b)$$

Since by (3.3b) the portfolio choice is the same for every agent, in order to clear the market for the risk-free asset, we must have  $1 - \theta = 0 \iff r = \mu - \gamma\sigma^2$ . Substituting  $r$  into the optimal consumption rule (3.3a), the marginal propensity to consume simplifies to  $m = \rho\varepsilon + (1 - \varepsilon)(\mu - \frac{\gamma\sigma^2}{2})$ . Substituting into the budget constraint (3.2), it follows that individual consumption  $c_{it}$  evolves according to the geometric Brownian motion

$$dc_{it}/c_{it} = gdt + \sigma dB_{it}, \quad (3.4)$$

where the expected growth rate of consumption is  $g = (\mu - \rho)\varepsilon - (1 - \varepsilon)\frac{\gamma\sigma^2}{2}$ .

Equation (3.4) shows that Gibrat (1931)'s law of proportionate growth holds for individual consumption. Therefore, if agents start with the same capital  $k_0$  and are infinitely lived, then the cross-sectional distribution of consumption is lognormal, where the mean of log consumption at time  $t$  is  $\log c_0 + (g - \frac{1}{2}\sigma^2)t$ ,  $c_0$  is initial consumption, the term  $-\frac{1}{2}\sigma^2$  comes from Ito's lemma, and the variance is  $\sigma^2 t$ .

Now we can derive the double power law in consumption with one twist to the model. Suppose that agents "die" at a constant Poisson rate  $\delta > 0$  and are reborn with initial capital  $k_0$ . We can interpret this situation as one in which entrepreneurs or dynasties go bankrupt at a constant rate and being replaced by new ones. The interpretation does not matter; what is important is that there is a mean-reverting force that prevents the distribution from becoming degenerate. Under the assumption of constant probability of birth/death, the problem of describing the size distribution of consumption becomes mechanistically equivalent to the model studied by Gabaix (2009). Consequently, the stationary cross-sectional density of consumption becomes

$$f_{dP}(x) = \begin{cases} \frac{\alpha\beta}{\alpha+\beta} c_0^\alpha x^{-\alpha-1}, & (x \geq c_0) \\ \frac{\alpha\beta}{\alpha+\beta} c_0^{-\beta} x^{\beta-1}, & (0 \leq x < c_0) \end{cases} \quad (3.5)$$

where  $c_0$  (the mode if  $\beta > 1$ ) is the consumption level corresponding to initial capital  $k_0$  and  $\alpha, \beta > 0$  are power law exponents of the upper and lower tails determined such that  $\zeta = -\alpha, \beta$  are solutions to the quadratic equation

$$\frac{\sigma^2}{2}\zeta^2 + \left( g - \frac{1}{2}\sigma^2 \right) \zeta - \delta = 0. \quad (3.6)$$

See Equation (20) in Gabaix (2009) for the derivation. The distribution (3.5) is known as the *double Pareto* distribution (Reed, 2001) and obeys the double power law with exponents  $(\alpha, \beta)$ .

The intuition for getting a stationary distribution is as follows. By Gibrat's law, the cross-sectional distribution of consumption *within* an age cohort is lognormal, and the variance increases linearly over time. But because agents die

and are reborn, there are exponentially fewer agents that live longer. These two effects balance with each other and generate the double Pareto distribution in the entire cross-section. We do not claim that this model is realistic, but we point out that (i) it is theoretically possible that consumption has fat tails, especially if (the permanent component of) consumption obeys Gibrat’s law, and (ii) consistent with Gibrat’s law, in actual data the consumption distribution within age cohorts is close to lognormal (Battistin et al., 2009) and the cross-sectional variance seems to increase linearly over time (Deaton and Paxson, 1994).

### 3.3 Evidence

In this section we study the tail behavior of the empirical consumption distribution.

#### 3.3.1 Data

We use the same data as the real, seasonally adjusted, quarterly household consumption data used in Kocherlakota and Pistaferri (2009) constructed from the Consumption Expenditure Survey (CEX). Their data is publicly available at the JPE website. Since households report the previous three month’s consumption but are surveyed in different months, we have 291 months of cross-sectional consumption data from December 1979 to February 2004. In one of the estimation exercises, we split the households into age cohorts. We obtain the age data from the NBER Consumer Expenditure Survey Family-Level Extracts webpage.<sup>2</sup>

#### 3.3.2 QQ plot

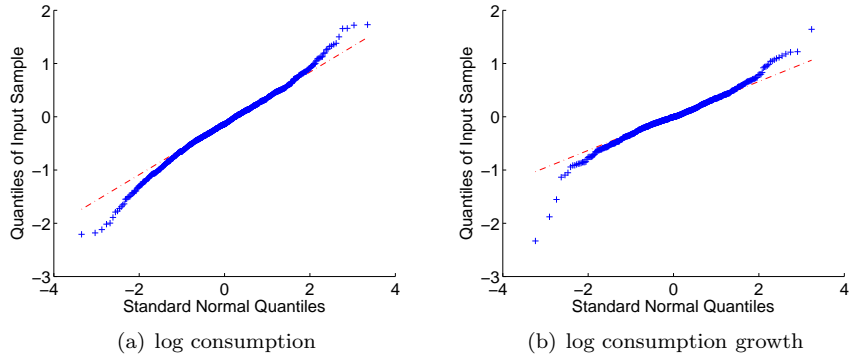
Figure 1 shows the QQ plot (quantile-quantile plot) of log consumption and log consumption growth against the standard normal distribution. If the variables are normally distributed, the points should lie around the 45 degree line. However, we can see that the points that are roughly two standard deviations from the mean deviate from the 45 degree line towards more extreme values. Therefore the QQ plot suggests that the distribution of consumption and consumption growth have fatter tails than lognormal. Although Figure 1 shows the results only for December 1980, other months look similar.

#### 3.3.3 Power law exponents

Since the model in Section 3.2 predicts that the tails obey the power law, we estimate the power law exponent of the upper and lower tails by maximum likelihood (Hill estimator) and perform the Kolmogorov test for goodness-of-fit.<sup>3</sup>

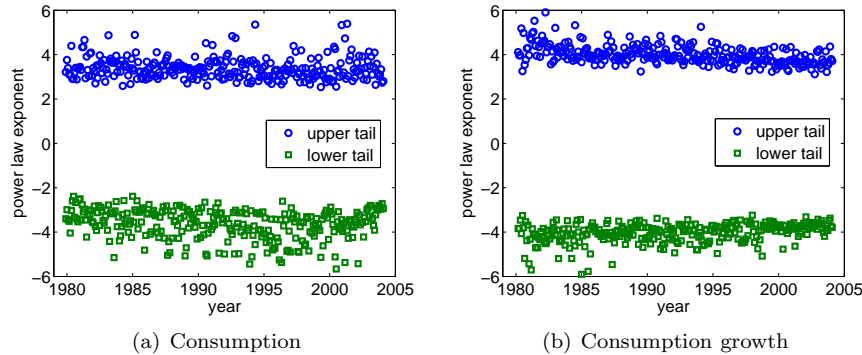
<sup>2</sup>[http://www.nber.org/data/ces\\_cbo.html](http://www.nber.org/data/ces_cbo.html)

<sup>3</sup>For this purpose we employ the Matlab files provided by Clauset et al. (2009), which can be downloaded from <http://tuvalu.santafe.edu/~aaronc/powerlaws/>. The file `plfit.m` estimates the power law exponent by maximum likelihood after choosing an appropriate cutoff value for the tail, and `plpva.m` performs the Kolmogorov test of goodness-of-fit by bootstrap (we choose the bootstrap repetition  $B = 500$ ). One caveat is that these authors define the power law exponent by  $\alpha' = \alpha + 1$ , so we need to subtract 1 from the output to convert to the usual definition. Also, in order to estimate the power law exponent of the lower tail, we need to input  $1/X$  (the reciprocal of consumption) instead of  $X$ .



**Figure 1.** Quantile-quantile plot against the normal distribution. December 1980.

Figure 2 shows the maximum likelihood estimates of the power law exponent of the upper and lower tails  $\alpha, \beta$  for each month. (We plot  $-\beta$  instead of  $\beta$  for visibility.) According to Figure 2, the power law exponents are around 4 for both tails and both consumption and consumption growth. The average across all months is  $(\bar{\alpha}, \bar{\beta}) = (3.38, 3.65)$  for consumption and  $(\bar{\alpha}, \bar{\beta}) = (3.99, 4.03)$  for consumption growth. The Kolmogorov test fails to reject the power law in consumption in 249 months out of 291 (86% of the time) for the upper tail and in 223 months (77% of the time) for the lower tail. With consumption growth, the power law is not rejected in 265 months out of 287 months (92% of the time) for both the upper and lower tails.



**Figure 2.** Maximum likelihood estimates of power law exponents.

### 3.3.4 Testing the existence of moments directly

Although documenting the power law in the cross-sectional consumption distribution is potentially interesting in its own right, in view of estimating Euler equations, whether a moment exists or not is more important. Fortunately, there is a simple bootstrap test for testing the existence of moments directly (Fedotenkov, 2013), which we explain briefly.

Suppose that the random variable  $X$  is nonnegative (consider  $|X|$  instead if  $X$  can be negative) and  $\{X_n\}_{n=1}^{\infty}$  are i.i.d. copies of  $X$ . If  $E[X^\eta] = \infty$ , then

the sample moment  $\frac{1}{N} \sum_{n=1}^N X_n^\eta$  tends to infinity as  $N \rightarrow \infty$ . Take a number  $M(N)$  such that  $M \rightarrow \infty$  and  $M/N \rightarrow 0$  as  $N \rightarrow \infty$  (so  $M$  tends to infinity at a slower rate than  $N$ , say  $M(N) = \sqrt{N}$ ), and let  $\{Y_m\}_{m=1}^\infty$  be i.i.d. copies of  $X$ . Then for  $0 < \xi < 1$ , the quantity

$$F = 1 \left\{ \frac{1}{M} \sum_{m=1}^M Y_m^\eta \geq \xi \frac{1}{N} \sum_{n=1}^N X_n^\eta \right\}$$

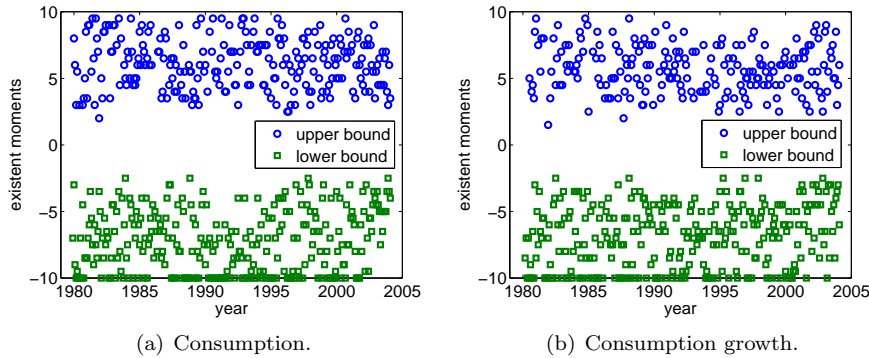
tends to zero almost surely as  $N \rightarrow \infty$ , where  $1\{\cdot\}$  denotes the indicator function (so  $F = 1$  if the inequality holds and  $F = 0$  otherwise). This limit holds because both  $\xi \frac{1}{N} \sum_{n=1}^N X_n^\eta$  and  $\frac{1}{M} \sum_{m=1}^M Y_m^\eta$  tend to infinity, but the former does so at a faster rate since  $N \gg M(N)$ . On the other hand, if  $E[X^\eta]$  is finite, then by the law of large numbers  $F$  tends to 1 almost surely because both sample means converge to the same population mean, but since  $0 < \xi < 1$  as  $N$  tends to infinity  $\xi \frac{1}{N} \sum_{n=1}^N X_n^\eta$  is almost surely smaller than  $\frac{1}{M} \sum_{m=1}^M Y_m^\eta$ .

Following this idea, Fedotenkov (2013) constructs a bootstrap test of moment existence as follows. Let  $\mathbf{x} = (x_1, \dots, x_N)$  be the data. First, we choose the bootstrap sample size  $M(N)$ , the parameter  $\xi$ , and bootstrap repetition  $B$  (Fedotenkov suggests taking  $M(N) = \lfloor \log N \rfloor$ ,  $\xi = 0.999$ , and  $B = 10,000$ ). Second, for each  $b = 1, \dots, B$ , we generate a bootstrap sample  $\mathbf{x}^b = (x_1^b, \dots, x_M^b)$  of size  $M$  drawn randomly with replacement from the data, and compute

$$F^b = 1 \left\{ \frac{1}{M} \sum_{m=1}^M (x_m^b)^\eta \geq \xi \frac{1}{N} \sum_{n=1}^N x_n^\eta \right\}.$$

Finally, the P value is defined by  $p = \frac{1}{B} \sum_{b=1}^B F^b$ .

Figure 3 shows the upper and lower bounds of the order of moments for which the existence is not rejected at significance level 0.05. The existence of moments starts to get rejected at around  $\eta = \pm 3$ , the same order of magnitude as the estimated power law exponents. The averages of the upper and lower bounds across all months are 6.73 and  $-7.16$  for consumption and 6.80, and  $-6.83$  for consumption growth. These numbers are slightly larger in magnitude than the estimated power law exponents (around 4).



**Figure 3.** Range of moment existence implied by the bootstrap test.

### 3.3.5 Tail thickness within age groups

So far we have presented evidence that household consumption and consumption growth have fat tails. Does this finding contradict to Battistin et al. (2009), who document that consumption is approximately lognormal (for which all moments exist)? The answer is no, because they look at the consumption distribution *within age cohorts*, not the entire cross-section. Since according to the model in Section 3.2 the double power law emerges from the birth/death (and consequently the exponential age distribution), we would expect that the cross-sectional consumption distribution is more lognormal within age cohorts than in the entire cross-section. To evaluate this conjecture, we perform the bootstrap test for moment existence for each age cohort, where we define the age of a household by the age of the oldest head of household.

The groups are household head age 30 or less, 31 to 40, 41 to 50, 51 to 60, and 60 or more. The range of moment existence is  $[-7.6, 10.3]$  for 30 or less,  $[-9.4, 8.9]$  for 31 to 40,  $[-10.5, 8.5]$  for 41 to 50,  $[-11.8, 8.5]$  for 51 to 60, and  $[-10.2, 6.9]$  for 61 or more. These ranges are wider than for the entire cross section.

## 4 GMM estimation and robustness

In this section we estimate the relative risk aversion coefficient  $\gamma$  using various asset pricing models and study the robustness of the performance of each model. In light of the results, we discuss the potential impacts of fat tails on GMM.

### 4.1 Data

As in Section 3, we use the real, seasonally adjusted consumption data in Kocherlakota and Pistaferri (2009) constructed from the CEX. Their dataset has monthly observations from December 1979 to February 2004, but each number corresponds to a household's consumption over the previous 3 months. So, while there are households for each month, no household appears in consecutive months. Therefore, even though we have a stochastic discount factor (SDF) and excess return realization for each month, the data for each month reflect a quarter of information, and the return series are 3 month moving averages. For example, the sample analog of the MU (marginal utility) SDF is defined by

$$\widehat{m}_t^{\text{MU}}(\gamma) = \frac{\frac{1}{I_t} \sum_{i=1}^{I_t} c_{it}^{-\gamma}}{\frac{1}{I_{t-3}} \sum_{i=1}^{I_{t-3}} c_{i,t-3}^{-\gamma}},$$

where  $I_t$  is the number of households at time  $t$  and  $c_{it}$  is the consumption of household  $i$  at time  $t$ . We have 288 SDF observations for RA (representative agent), MU (marginal utility), and PIPO (private information with Pareto optimality) and 287 for IMRS (intertemporal marginal rate of substitution; we lose one quarter for IMRS because household IDs were reset in 1986). In total, we have 410,788 consumption data points and 270,428 consumption growth data points. There are fewer consumption growth data points because many households participate the survey for only one quarter, in which case we have no data on consumption growth. See Kocherlakota and Pistaferri (2009) for further details on the construction of real consumption and the U.S. equity premium.

## 4.2 GMM estimation

For any stochastic discount factor  $j \in \{\text{RA}, \text{IMRS}, \text{MU}, \text{PIPO}\}$ , let

$$g_T^j(\gamma) = \frac{1}{T} \sum_{t=1}^T \widehat{m}_t^j(\gamma)(R_t^s - R_t^b)$$

be the sample average of the pricing error for the equity premium, where  $T$  is the number of observation for SDF  $j$ ,  $R_t^s$  is the stock market return, and  $R_t^b$  is the Treasury bill rate. The GMM estimator of the relative risk aversion coefficient  $\gamma$  and the pricing error are

$$\begin{aligned} \widehat{\gamma}^j &= \arg \min_{\gamma} T g_T^j(\gamma)^2, \\ e^j &= g_T^j(\widehat{\gamma}^j) = \frac{1}{T} \sum_{t=1}^T \widehat{m}_t^j(\widehat{\gamma}^j)(R_t^s - R_t^b). \end{aligned}$$

For standard errors, we report both the Newey-West standard errors (with truncation parameter equal to 4) and bootstrap ones. The Newey-West standard errors account for the sampling error in the time series but abstract from uncertainty regarding cross-sectional moments of consumption. The bootstrap standard errors are based on the stationary bootstrap of Politis and Romano (1994), which also account for the sampling error in the cross-section as well as the time series. We sample with replacement from the original data to generate  $B$  bootstrap samples, indexed by  $b = 1, \dots, B$ . Each is of length  $T$  and has statistical properties like the original sample. Each bootstrap sample yields risk aversion estimate  $\widehat{\gamma}_b^j$  and pricing error  $e_b^j$ . The bootstrap standard error is the sample standard error of  $\left\{ \widehat{\gamma}_b^j \right\}_{b=1}^B$ . The explicit procedure for generating each sample  $b$  is as follows:

1. For each  $t \in \mathbf{T} = \{1, \dots, T\}$ , draw with replacement  $I_t$  observations from  $\{c_{it}\}_{i=1}^{I_t}$ , yielding  $\{c_{it}^b\}_{i=1}^{I_t}$ .
2. Let  $M$  be the average block length and set  $p = 1/M$ . (We choose  $M = \sqrt{T}$ .) Draw  $\tau_1^b$  uniformly from  $\mathbf{T}$ . For  $s = 2, \dots, T$ , with probability  $1 - p$  set  $\tau_s^b = \tau_{s-1}^b + 1$  modulo  $T$  (hence  $\tau_s^b = 1$  if  $\tau_{s-1}^b = T$ ), and with probability  $p$  draw  $\tau_s^b$  uniformly from  $\mathbf{T}$ .
3. The bootstrap sample  $b$  consists of all  $\widetilde{c}_{is}^b$ ,  $s = 1, \dots, T$ , where we define  $\widetilde{c}_{is}^b = c_{i, \tau_s^b}^b$  for  $i = 1, \dots, I_{\tau_s^b}$ .

The process for bootstrapping consumption growth and asset returns is analogous. The one caveat concerns the calculation of SDF  $j \in \{\text{RA}, \text{MU}, \text{PIPO}\}$ . Consider MU for example. We use

$$\widehat{m}_s^{\text{MU}, b}(\gamma) = \frac{\frac{1}{I_{\tau_s^b}} \sum_{i=1}^{I_{\tau_s^b}} (\widetilde{c}_{is}^b)^{-\gamma}}{\frac{1}{I_{\tau_s^b-3}} \sum_{i=1}^{I_{\tau_s^b-3}} (c_{i, \tau_s^b-3}^b)^{-\gamma}}.$$

That is, the bootstrap time  $s$  SDF is formed from actual time  $\tau_s^b$  and  $\tau_s^b - 3$  data, not actual time  $\tau_s^b$  and  $\tau_{s-3}^b$  data in order to preserve the statistical properties of the SDF. Below, we use  $B = 500$  bootstrap replications.

### 4.3 Results and robustness

#### 4.3.1 Estimation with full sample

The column “Full Sample” in Table 2 shows the estimation results. The first and second numbers in parentheses are the Newey-West and bootstrap standard errors, respectively. For RA and PIPO, the bootstrap standard errors (which account for cross-sectional sampling error) are similar to the Newey-West ones. This similarity does not hold for MU and IMRS. However, it is not clear how to interpret these Newey-West numbers, for in each case  $\gamma$  is exactly identified but the pricing error is away from 0 (as returns are quarterly, 0.019 is essentially the entire equity premium). Another reason for the large standard errors in MU and IMRS might be again the fat tails. For example, even if the pricing error  $\widehat{m}_t^j(\gamma)(R_t^s - R_t^b)$  has a finite first moment, it may not have a finite second moment, in which case we cannot apply the standard asymptotic theory of the GMM estimator.

$\widehat{\gamma}^{\text{IMRS}}$  is close to zero. One explanation is that since the exact IMRS SDF is the weighted average of the  $-\gamma$ -th power of each household’s consumption growth, whenever  $\gamma$  is large the SDF will be huge because there are always households with consumption growth much smaller than 1. Therefore the GMM criterion may be a huge number when  $\gamma$  is large. In this way, small consumption growth observations may drive  $\gamma$  toward zero.

**Table 2.** GMM estimation of relative risk aversion (RRA)  $\gamma$  and pricing errors  $e^j$ . The first and second numbers in parentheses are the Newey-West and bootstrap standard errors.

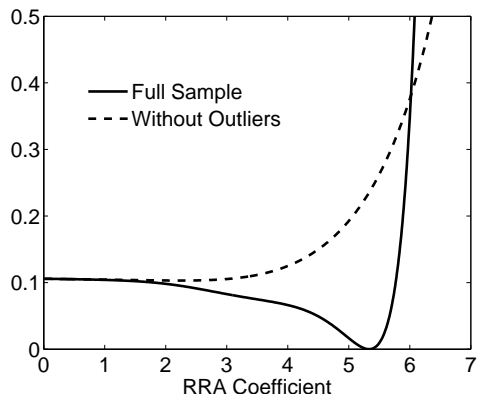
| Model | Full Sample                 |               | Without Outliers            |               |
|-------|-----------------------------|---------------|-----------------------------|---------------|
|       | RRA ( $\gamma$ )            | Pricing error | RRA ( $\gamma$ )            | Pricing error |
| RA    | 53.26<br>(29.41)<br>(20.19) | 0.000         | 53.10<br>(30.85)<br>(21.24) | -0.000        |
| IMRS  | 0.03<br>(1035)<br>(0.08)    | 0.019         | 0.03<br>(1297)<br>(0.21)    | 0.019         |
| MU    | 1.52<br>(5698)<br>(0.90)    | 0.019         | 2.51<br>(9960)<br>(1.86)    | 0.019         |
| PIPO  | 5.33<br>(1.42)<br>(1.98)    | 0.000         | 2.23<br>(8010)<br>(1.68)    | 0.019         |

#### 4.3.2 Estimation without outliers

The column “Without Outliers” in Table 2 shows the estimation results when we drop a small number of outliers relative to the total number of data points. Specifically, we drop the top 100 and bottom 100 consumption observations from the entire sample. For IMRS, we also drop the top 100 and bottom 100 consumption growth observations. As there are 410,788 data points, for consumption levels the points we drop account for less than 0.05% of the entire sample. Note that these outliers are spread roughly uniformly across the quar-

ters, so on average we are dropping less than 1 observation point per quarter since there are 288 quarters.

We see that the results for the RA SDF, which should not be affected by the nonexistence of higher moments, barely change. We continue to reject MU and IMRS. PIPO, however, no longer explains the equity premium. Figure 4 below shows the GMM criterion for PIPO as a function of  $\gamma$ , with and without the outliers. Just a few outliers generate the trough at 5.33.



**Figure 4.** PIPO GMM criterion with and without largest and smallest 100 consumption outliers out of 410,788.

### 4.3.3 Examination of bootstrap samples

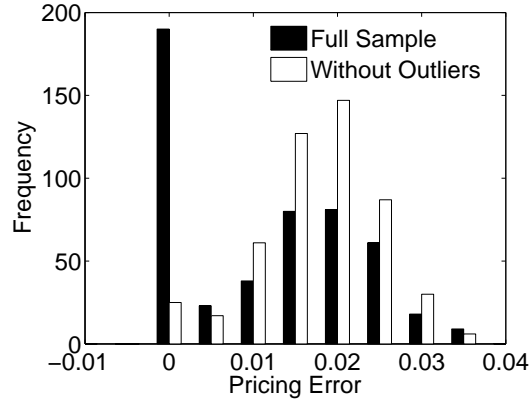
We also analyze the bootstrap distributions for the pricing error and  $\gamma$  estimate. Figure 5 displays histograms of  $e_b^{\text{PIPO}}$ , with and without outliers. We see that when we bootstrap with all data, there is a mass of pricing errors at 0. Without outliers, the pricing error bootstrap distribution is centered around  $e^{\text{PIPO}}$ , as it should be, with much less mass at zero. Although we do not have an explanation of the appearance and disappearance of the bimodal pricing error, fat tails may well be the cause. For example, Fiorio et al. (2010) show that the limit distribution of the  $t$ -statistic for the mean can become bimodal when the underlying distribution has fat tails.

Finally, Figure 6 shows a scatter plot of the bootstrap estimates  $\hat{\gamma}_b^j$  and pricing errors  $e_b^j$ . There is an inverse relationship between the pricing error and the  $\gamma$  estimate. Indeed, most of the zero pricing errors correspond to  $\gamma$  estimates in the moment nonexistence range ( $> 4$ ); when the pricing error is greater than 0.01, the corresponding  $\gamma$  estimate tends to be less than 3.

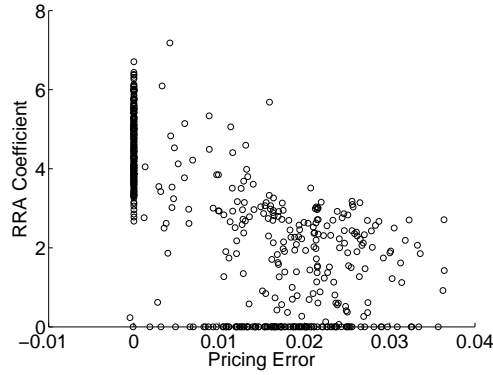
We take this collection of observations as evidence that the fat tails of the consumption distribution may aid mechanically in zeroing the pricing error. At least, such CEX-based asset pricing exercises seem quite sensitive to outliers.<sup>4</sup>

<sup>4</sup>Of course, one may argue that the reason why the PIPO model fails to explain the equity premium without outliers is that the rich plays an important role in asset pricing. However, this interpretation does not seem plausible because if it were the case the histogram of the bootstrapped pricing errors (with the entire sample) should be centered around 0, while in fact it is bimodal as in Figure 5.





**Figure 5.** Histogram of bootstrapped PIPO pricing errors with and without largest and smallest 100 consumption outliers out of 410,788.



**Figure 6.** Scatter plot of bootstrapped PIPO  $\gamma$  estimates and pricing errors.

## 5 Potential solution: conditioning on age

How should one estimate and test Euler equations if consumption has fat tails? Since dropping outliers mitigated the fat tail issue, one may think that trimming the tails is the solution. However, this practice is problematic since consumption is an endogenous variable. To see this for the IMRS SDF, observe that the theory predicts the (unconditional) Euler equation

$$1 = E[\beta g_{t+1}^{-\gamma} R_{t+1}] \quad (5.1)$$

where  $g_{t+1} = c_{t+1}/c_t$  is consumption growth. Suppose, for example, that the researcher trims the tails of consumption growth by dropping observations outside the range  $[g, \bar{g}]$ . Then the researcher is in fact testing the conditional moment restriction

$$1 = E[\beta g_{t+1}^{-\gamma} R_{t+1} \mid g \leq g_{t+1} \leq \bar{g}], \quad (5.2)$$

which is different from (5.1). Note that even if the model is correct (*i.e.*, (5.1) holds), the conditional moment restriction (5.2) is almost always false for generic thresholds  $(g, \bar{g})$ .

One solution is to find an exogenous conditioning variable such that the conditional consumption distribution does not have fat tails. When we tested

the consumption double power law conjecture in Section 3.3, for each quarter  $t$  we divided the cross-section into five age cohorts, 30 years or younger, 31 to 40, 41 to 50, 51 to 60, and older than 60. Call these  $H_{t,1}, \dots, H_{t,5}$ . We found that at each  $t$ , within cohort the consumption distribution is approximately lognormal (see also Battistin et al. (2009)). At least, more moments exist within cohorts than for the entire cross-section. Furthermore, in the continuous-time limit we explore in our model, the within age group distribution is precisely lognormal. With this pattern in mind, we perform an overidentified GMM exercise that (i) seems less susceptible to the nonexistent moment issue and (ii) allows for overidentifying tests of the different models.

Specifically, we exploit the fact that the Euler equation aggregation in Section 2.1 that gave us the SDFs also works within a particular age cohort because age is an exogenous variable. That is, instead of averaging across all agents, we can average across a particular age group. For example, we can form the  $H_{t,5}$  ( $> 60$ ) MU SDF by

$$\widehat{m}_t^{\text{MU},5}(\gamma) = \frac{\frac{1}{|H_{t,5}|} \sum_{i \in H_{t,5}} c_{it}^{-\gamma}}{\frac{1}{|H_{t-3,5}|} \sum_{i \in H_{t-3,5}} c_{i,t-3}^{-\gamma}},$$

where  $|H_{t,5}|$  is the number of households in group  $H_{t,5}$ .

For any  $j \in \{\text{RA}, \text{IMRS}, \text{MU}, \text{PIPO}\}$ , let  $\widehat{\mathbf{m}}_t^j(\gamma) = (\widehat{m}_t^{j,1}(\gamma), \dots, \widehat{m}_t^{j,5}(\gamma))'$  be the vector of SDFs and

$$G_T^j(\gamma) = \frac{1}{T} \sum_{t=1}^T \widehat{\mathbf{m}}_t^j(\gamma) (R_t^s - R_t^b)$$

be the vector of pricing errors. The overidentified GMM estimator of  $\gamma$  is

$$\widehat{\gamma}^j = \arg \min_{\gamma} T G_T^j(\gamma)' W G_T^j(\gamma),$$

where  $W$  is the weighting matrix. (We always use the identity matrix as the weighting matrix.)

We calculate standard errors via the above bootstrap procedure because the Newey-West standard errors may be misleading according to the results of Table 2. Furthermore, for each SDF we bootstrap a P value for the null hypothesis that the pricing error is 0 (that is, that the model is correct). The following is a description of the calculation of these P values:

1. Dropping the SDF superscript, let  $G_{T,b}(\widehat{\gamma}_b)$  be the vector of pricing errors corresponding to bootstrap sample  $b$ .
2. For each bootstrap sample  $b$ , define

$$J_{T,b} = T (G_{T,b}(\widehat{\gamma}_b) - G_T(\widehat{\gamma}))' W (G_{T,b}(\widehat{\gamma}_b) - G_T(\widehat{\gamma})).$$

Also define the minimized sample criterion  $J_T = T G_T(\widehat{\gamma})' W G_T(\widehat{\gamma})$ .

3. Calculate the P value by  $p = \frac{1}{B} \sum_{b=1}^B 1 \{J_{T,b} \geq J_T\}$ .

Why should this procedure work? The idea of the bootstrap is that the empirical distribution of  $G_{T,b}$  around  $G_T$  approximates the distribution of  $G_T$

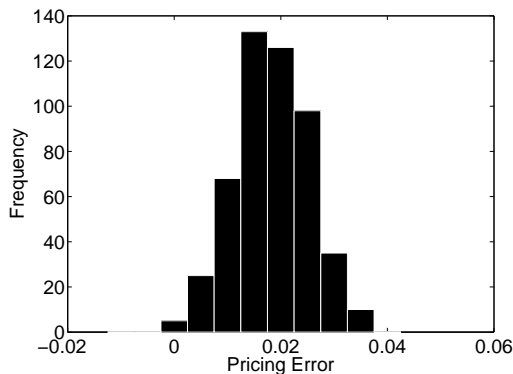
around  $G_\infty$ , which is 0 under the null. It follows that under the null the empirical distribution of  $J_{T,b}$  approximates the distribution of  $J_T$ . Finally, if the null fails and  $G_T$  converges to something different from 0, then  $J_T$  is not properly centered and will diverge as  $T \rightarrow \infty$ .

Table 3 presents the age cohort GMM  $\gamma$  estimates and the bootstrapped P values. We see that with the age cohort method, the RA, MU, and PIPO  $\gamma$  estimates are all between 1 and 3, well within the moment existence range. The IMRS estimate, as before, is around 0. The standard errors for the former three SDFs are, respectively, 1.68, 0.63, and 0.88, meaning the risk aversion estimates of these models are statistically close. Moreover, the overidentifying tests reject all of the models we consider at the 1% significance level.

**Table 3.** Age cohort GMM estimation of relative risk aversion (RRA)  $\gamma$  and P value of overidentifying tests. Numbers in parentheses are bootstrapped standard errors.

| Model | RRA ( $\gamma$ ) | P value |
|-------|------------------|---------|
| RA    | 2.62<br>(1.68)   | 0.00    |
| IMRS  | 0.04<br>(0.10)   | 0.00    |
| MU    | 1.22<br>(0.63)   | 0.00    |
| PIPO  | 1.88<br>(0.88)   | 0.00    |

The rejection of IMRS and MU is not surprising, since these models do not zero the pricing error even in the single equation case. The fact that the  $\gamma$  estimate for RA drops from about 50 to 2 by dividing the population into age groups suggests that raising aggregate consumption to a high power is problematic. With respect to PIPO, the model rejection and low  $\gamma$  estimate are further evidence that the power law may interfere with estimation and model selection. One more piece of evidence is Figure 7, which is a histogram of the average pricing error,  $(\mathbf{1}'G_T(\hat{\gamma}))/5$ , across bootstrap samples. As when we drop outliers (compare to Figure 5), there is no spike at 0.



**Figure 7.** Histogram of bootstrapped PIPO pricing errors with age cohort GMM estimation.

While the approximate lognormality within age groups was our original jus-

tification for age cohort GMM, we conjecture that there is another reason this procedure helps mitigate spurious non-rejection and inconsistency from fat tails. As we have argued, sample analogs of nonexistent moments seem to generate spurious and relatively sharp troughs in GMM criteria. While these troughs appear in the nonexistence range, their precise location seems random and sample dependent. Therefore, even if a subset of the moment conditions in age cohort GMM exhibits such troughs, they may be impossible to exploit when there is only one unknown parameter ( $\gamma$ ): a high  $\gamma$  may zero the pricing error for one cohort, only to blow up the other pricing errors, which have different or no spurious troughs. The minimum of the joint criterion must then lie near the true  $\gamma$ , which uniquely zeroes *all* moment conditions if the model is true. With sufficiently many moment conditions, age cohort GMM may be robust to fat tails even if one cohort still exhibits the power law. As we cannot definitively rule out fat tails within cohorts, this interpretation improves the reliability of the results in Table 3. Indeed, with a finite number of cohorts, at least one age group must exhibit fat tails if the whole distribution does, since the moment of the whole distribution is the weighted average of those of age groups. This property is not true, however, when there are an infinite number of cohorts (as in our continuous-time model).

## References

- 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.
- 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.
- 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.
- Aaron Clauset, Cosma Rohilla Shalizi, and Mark E. J. Newman. Power-law distributions in empirical data. *SIAM Review*, 51(4):661–703, 2009. doi:10.1137/070710111.
- 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.
- Angus Deaton and Christina Paxson. Intertemporal choice and inequality. *Journal of Political Economy*, 102(3):437–467, June 1994.
- Igor Fedotenkov. A bootstrap method to test for the existence of finite moments. *Journal of Nonparametric Statistics*, 25(2):315–322, 2013. doi:10.1080/10485252.2012.752487.

- William Feller. A limit theorem for random variables with infinite moments. *American Journal of Mathematics*, 68(2):257–262, April 1946.
- 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.
- Robert Gibrat. *Les Inégalités économiques*. Librairie du Recueil Sirey, Paris, 1931.
- Kristian Giesen, Arndt Zimmermann, and Jens Suedekum. The size distribution across all cities—double Pareto lognormal strikes. *Journal of Urban Economics*, 68(2):129–137, September 2010. doi:10.1016/j.jue.2010.03.007.
- 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.
- 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.
- 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.
- Benôit Mandelbrot. The Pareto-Lévy law and the distribution of income. *International Economic Review*, 1(2):79–106, May 1960. doi:10.2307/2525289.
- Robert C. Merton. Optimum consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory*, 3(4):373–413, December 1971. doi:10.1016/0022-0531(71)90038-X.
- Robert A. Miller. Estimating models of dynamic optimization with microeconomic data. In M. Hashem Pesaran and Peter Schmidt, editors, *Handbook of Applied Econometrics*, volume 2, chapter 6, pages 246–299. Blackwell Publishing, 1999.
- Vilfredo Pareto. *Cours d'économie politique*. F. Rouge, Lausanne, 1896.
- 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.

Alexis Akira Toda. The double power law in income distribution: Explanations and evidence. *Journal of Economic Behavior and Organization*, 84(1):364–381, September 2012. doi:10.1016/j.jebo.2012.04.012.

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.