The Double Power Law in Consumption and Implications for Estimating Asset Pricing Models

Alexis Akira Toda[1] Kieran Walsh[2]

[1]UCSD

[2]Yale University

December 12, 2013

U.S. household log consumption is bell-shaped...

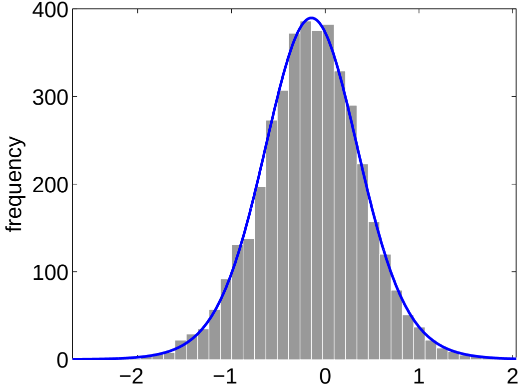


Figure: 1985 Q1

. . . but has heavier tails than normal.

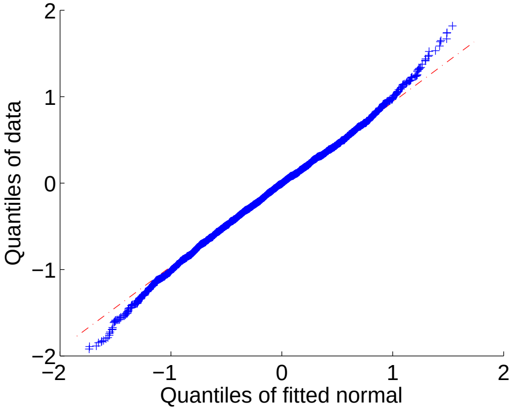


Figure: 1985 Q1

Questions

1. How is consumption distributed?
2. Why does consumption exhibit heavy (power law) tails?
3. What are the practical implications of power law, especially for estimating heterogeneous-agent, consumption-based asset pricing models (cCAPM)?
4. How can we circumvent the power law issue in estimation?

Contributions

1. Establish double power law (Toda, 2012a) in U.S. cross-sectional household consumption data, with power law exponent ≈ 4 in both tails.
2. Holding the data sample constant, reestimate and compare all heterogeneous-agent asset pricing models of the literature.
   - Provide evidence suggesting power law tails cause GMM spuriousness. (Main tool: bootstrap samples.)
   - Propose power law-robust estimation methods but reject all existing models. (Exploit within cohort log-normality.)
3. Illustrate GMM failure through Monte Carlo studies with simulated data from the GE model of Toda (2012b), which generates the consumption power law.

Literature

Power law Pareto (1896), Mandelbrot (1960, 1961, 1963), Gabaix (1999, 2009), Reed (2001, 2003), Toda (2011, 2012a, 2012b)

Heterogeneous-agent cCAPM Constantinides & Duffie (1996), Brav Constantinides, & Geczy (2002), Cogley (2002), Balduzzi & Yao (2007), Kocherlakota & Pistaferri (2009), Ludvigson (2013)

Double power law

- Nonnegative random variable X obeys
  - power law (in upper tail) if $P(X > x) \sim x^{-\alpha}$ as $x \to \infty$,
  - power law (in lower tail) if $P(X < x) \sim x^{\beta}$ as $x \to 0$,
  - double power law if power law holds in both tails.
- $\alpha, \beta > 0$: power law exponents.
- Canonical example is double Pareto distribution (Reed, 2001):

$$f(x) = \begin{cases} \frac{\alpha\beta}{\alpha+\beta} \frac{1}{M} \left(\frac{x}{M}\right)^{-\alpha-1}, & (x > M) \\ \frac{\alpha\beta}{\alpha+\beta} \frac{1}{M} \left(\frac{x}{M}\right)^{\beta-1}, & (x \leq M) \end{cases}$$

  where M: mode.
- Product of independent double Pareto and lognormal variables = double Pareto-lognormal (dPlN) (Reed, 2003).

Laplace and normal-Laplace (NL) distributions

- Log of double Pareto is Laplace distribution:

$$f(x) = \begin{cases} \frac{\alpha\beta}{\alpha+\beta} e^{-\alpha(x-m)}, & (x > m) \\ \frac{\alpha\beta}{\alpha+\beta} e^{\beta(x-m)}, & (x \le m) \end{cases}$$

  where $m = \log M$: mode.

- If X: normal and Y: Laplace, then $X + Y$: normal-Laplace.

Z	$\log Z$
lognormal	normal
double Pareto	Laplace
dPIN	normal-Laplace

Theoretical emergence of double power law

GE model of Toda (2012b):

- With multiplicative (aggregate & idiosyncratic) shocks and homothetic preferences, an agent's consumption is proportional to wealth $\Longrightarrow$ Gibrat's law. Model
- Geometric sums of independent random variables are approximately Laplace. Limit theorem
- Hence if agents (households/dynasties) die at a constant Poisson rate and are reborn (so age distribution is geometric), the cross-sectional distribution of log wealth is approximately Laplace, independent of the return distribution.
- Wealth and consumption distributions are double Pareto, because Laplace is log of double Pareto.

Data

- Kocherlakota and Pistaferri (2009):
  - Available on JPE website.
  - Consumer Expenditure Survey from December 1979 to February 2004.
  - Nondurable consumption, seasonally adjusted, deflated by CPI.
- NBER website CEX extracts:
  - Age of head of household (for forming cohorts).

Histogram of log consumption and normal-Laplace density

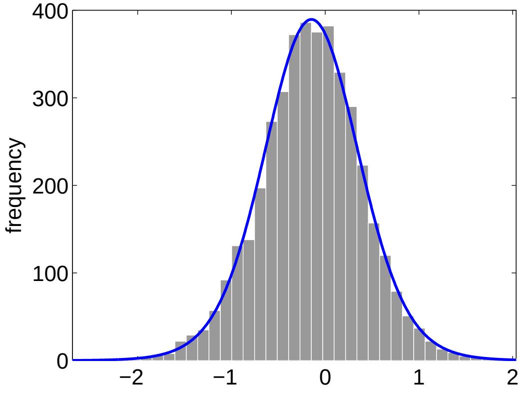


Figure: 1985 Q1

QQ plot (normal distribution)

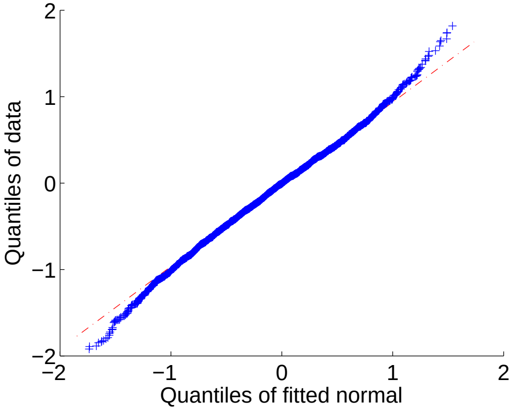


Figure: 1985 Q1

QQ plot (normal-Laplace distribution)

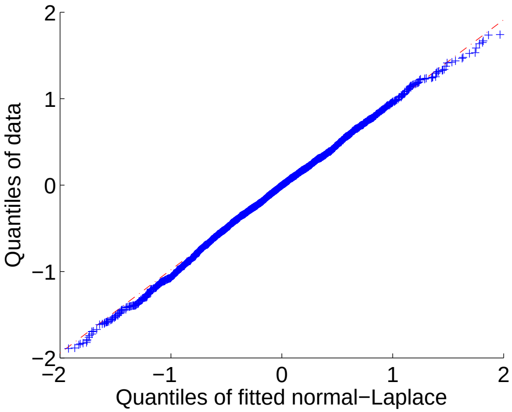


Figure: 1985 Q1

Power law exponent

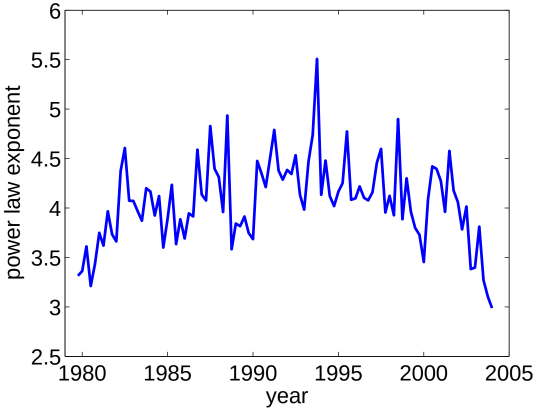


Figure: Power law exponent

Evidence for consumption

- Kolmogorov-Smirnov test fails to reject dPIN in 79 out of 98 quarters, while rejects lognormal in 73 quarters at significance level 0.05.
- Anderson-Darling test (puts more weight on tails) fails to reject dPIN in 64 quarters, while rejects lognormal in 92 quarters.
- Likelihood ratio test rejects lognormal against dPIN in all but 1 quarter.
- Power law exponent ≈ 4 and in the range $[3.0, 5.5]$.
- Double power law disappears within age cohort (c.f., Battistin, Blundel, & Lewbel, 2009).
  $\Rightarrow$ Suggests importance of age for generating power law, as predicted by model.

Histogram of log consumption growth

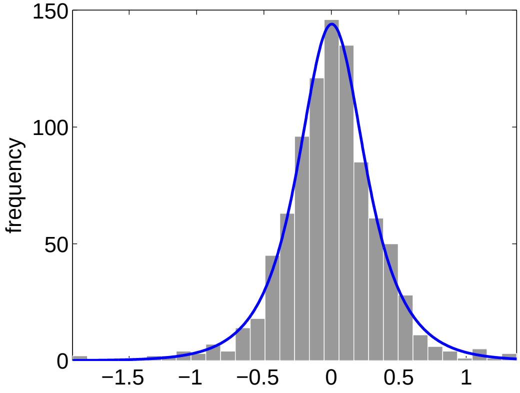


Figure: March 1985

Power law exponent

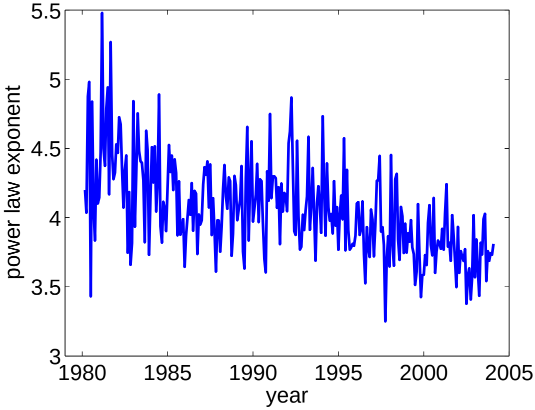


Figure: Power law exponent

Evidence for consumption growth

- Kolmogorov-Smirnov test fails to reject dPIN in 265 out of 291 months, while rejects lognormal in all but 3 months at significance level 0.05.
- Anderson-Darling test (puts more weight on tails) fails to reject dPIN in 260 months, while rejects lognormal in every single month.
- Likelihood ratio test rejects lognormal against dPIN in every single month.
- Power law exponent ≈ 4 and in the range $[3.0, 5.5]$.

Baseline Model

- Agents $i = 1, 2, \dots, I$.
- Assume agents have identical CRRA utility function

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

where β: discount factor, γ: relative risk aversion.

- Assume agents face budget constraints but no financing frictions.

Asset pricing in complete markets

- Euler equation
$$c_{it}^{-\gamma} = \mathrm{E}_{it}[\beta c_{i,t+1}^{-\gamma} R_{t+1}],$$
where i: household index, R_{t+1}: any gross asset return.
- Stochastic discount factor or pricing kernel: $\beta(c_{i,t+1}/c_{it})^{-\gamma}$.
- With complete markets, can substitute aggregate consumption.
- Equity premium puzzle: aggregate consumption and equity premium data are not consistent with this model (Hansen & Singleton, 1982, 1983; Mehra & Prescott, 1985).

Asset pricing in incomplete markets

- Aggregate consumption is relevant only with complete markets $\Longrightarrow$ incomplete markets models and consumption panel data (CEX).
- Euler equation
$$c_{it}^{-\gamma} = \mathrm{E}_{it}[\beta c_{i,t+1}^{-\gamma} R_{t+1}]$$
still valid in incomplete markets.
- What if c_{it}'s are measured poorly or we only see agent i for a few t's?

Aggregation 1

- Rewrite Euler equation as $1 = \mathrm{E}_{it}[\beta(c_{i,t+1}/c_{it})^{-\gamma} R_{t+1}]$.
- Taking unconditional expectations with respect to i,

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

- Ignoring β,

$$\widehat{m}_{t+1}^{\mathrm{IMRS}} = \frac{1}{I} \sum_{i=1}^{I} \left(\frac{c_{i,t+1}}{c_{it}} \right)^{-\gamma} \approx \mathrm{E}_{t+1}[(c_{i,t+1}/c_{it})^{-\gamma}]$$

  is a valid stochastic discount factor (SDF)
  (Brav, Constantinides, & Geczy, 2002; Cogley, 2002).
  (IMRS = Intertemporal Marginal Rate of Substitution)

Aggregation 2

- Averaging Euler equation $c_{it}^{-\gamma} = \mathrm{E}_{it}[\beta c_{i,t+1}^{-\gamma} R_{t+1}]$ directly,

$$\widehat{m}_{t+1}^{\mathrm{MU}} = \widehat{C}_{-\gamma,t+1}/\widehat{C}_{-\gamma,t}$$

  is a valid SDF (Balduzzi & Yao, 2007), where $\widehat{C}_{\eta,t} := \frac{1}{I}\sum_{i=1}^{I} c_{it}^{\eta}$ is η-th cross-sectional sample moment of consumption. (MU = Marginal Utility)
- Kocherlakota & Pistaferri (2009) start with inverse Euler equation and derive

$$\widehat{m}_{t+1}^{\mathrm{PIPO}} = \widehat{C}_{\gamma,t}/\widehat{C}_{\gamma,t+1}.$$

  (PIPO = Private Information with Pareto Optimality)

Estimation and Fit

- For $j \in \{\mathrm{RA}, \mathrm{IMRS}, \mathrm{MU}, \mathrm{PIPO}\}$, acquire/clean CEX data, and let

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

$$\widehat{\gamma}^j = \arg\min_{\gamma} T\left(PE_T^j(\gamma)\right)^2,$$

where $\widehat{m}_t^j(\gamma)$ is SDF, R_t^s: stock return, R_t^b: bond return.

- The pricing error is $PE_T^j\left(\widehat{\gamma}^j\right)$.
- Reject the model if
  - $PE_T^j\left(\widehat{\gamma}^j\right) \neq 0$ when exactly identified,
  - high J-statistic when over-identified.

RRA γ and tests of SDFs in literature

Paper	Sample	IMRS	MU	PIPO
BCG (2002)	1982–1996	✓3.5		
Cogley (2002)	1980–1994	X		
BY (2007)	1982–1995	Q: ✓5 M: X	✓10	
KP (2009)	1980–2004		X	✓5

- In order for GMM estimation to be consistent, need
  - $-\gamma$-th moment of consumption growth for IMRS,
  - $-\gamma$-th moment of consumption for MU,
  - γ-th moment of consumption for PIPO.
- But all γ estimates in moment non-existent region!

Ludvigson (2013)

"The mixed results seem to depend sensitively on a number of factors, including the sample, the empirical design, on the method for handling and modeling measurement error, the form of cross-sectional aggregation of Euler equations across heterogeneous agents, and the implementation, if any, of linear approximation of the pricing kernel. A tedious but productive task for future work will be to carefully control for all of these factors in a single empirical study, so that researchers may better assess whether the household consumption heterogeneity we can measure in the data has the characteristics needed to explain asset return data."

Estimation of asset pricing models

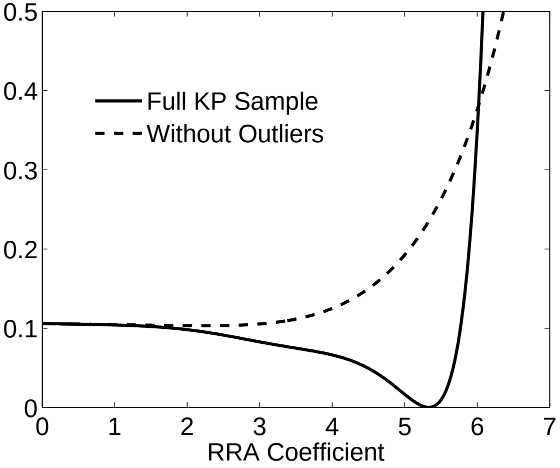
Estimate RA (representative agent), IMRS ?, MU ?, PIPO ? model by GMM for

1. single equation with full sample,
2. single equation without top and bottom 100 consumption data points out of 410,708 (dropping less than 0.05% of sample),
3. overidentified model by dividing sample into age cohorts (household head age 30 or less, 30s, 40s, 50s, 60 or more).

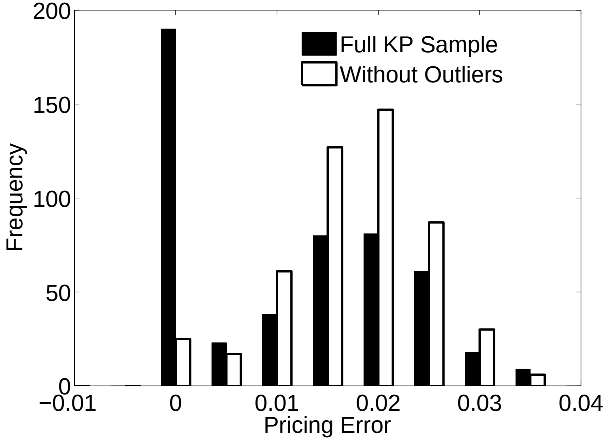
Results with and without top and bottom 100 points

	Full KP sample		Without Outliers	
Model	RRA (γ)	Pricing error	RRA (γ)	Pricing error
RA	53.26 (29.41) (20.19)	0.000	53.10 (30.85) (21.24)	-0.000
IMRS	0.03 (1035) (0.08)	0.019	0.03 (1297) (0.21)	0.019
MU	1.52 (5698) (0.90)	0.019	2.51 (9960) (1.86)	0.019
PIPO	5.33 (1.42) (1.98)	0.000	2.23 (8010) (1.68)	0.019

GMM criterion with and without top and bottom 100

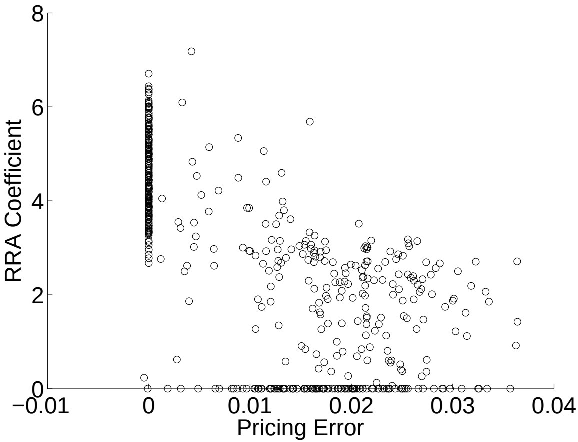


Histogram of bootstrapped PIPO pricing errors



Tool: stationary bootstrap (Politis & Romano, 1994) ?

Scatter plot of bootstrapped PIPO RRA estimates and pricing errors



Robust estimation using age data

- According to theory, power law emerges from geometric age distribution. ▸ ?
- In fact, power law disappears within age cohorts (Battistin, Blundel, & Lewbel, 2009).
- Dividing sample into age cohorts, we can estimate and test overidentified model without power law issue.

$$PE_T^j(\gamma) = \frac{1}{T}\sum_{t=1}^{T} \begin{bmatrix} \widehat{m}_t^{j,1}(\gamma) \\ \vdots \\ \widehat{m}_t^{j,5}(\gamma) \end{bmatrix} \left(R_t^s - R_t^b\right)$$

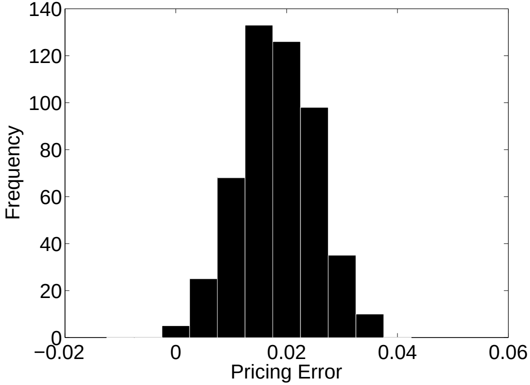
$$\widehat{\gamma}^j = \arg\min_{\gamma} T\left(PE_T^j(\gamma)\right)' W \left(PE_T^j(\gamma)\right)$$

Robust GMM estimation of RRA γ and P value

Model	RRA (γ)	P value
RA	2.62 (1.68)	0.00
IMRS	0.04 (0.10)	0.00
MU	1.22 (0.63)	0.00
PIPO	1.88 (0.88)	0.00

- γ estimates reasonable, but all models rejected.

Histogram of bootstrapped PIPO pricing errors with robust GMM estimation



Spurious peak at 0 disappears! compare

Using simulated data

- So far, our evidence is indirect since assume model is true.
- Here we construct an artificial economy from a model based on Toda (2012b) details and estimate parameters using both true and false data.
- Agents $i = 1, \dots, I$ have identical additive CRRA preferences and two linear technologies (stock market & private equity).
- Both technologies subject to aggregate shock, private equity subject to idiosyncratic shock.
- MU SDF ? is valid.

Calibration

Table: Parameters

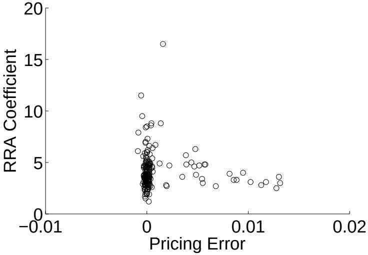
β	0.96		
γ	4		
δ	1/75		
μ_s	0.07	σ_s	0.15
μ_p	0.07	σ_p	0.1
ρ	0.5	σ_i	0.15
I	4000	σ_0	0.5
T	300	σ_ϵ	0.25

Table: Implied values

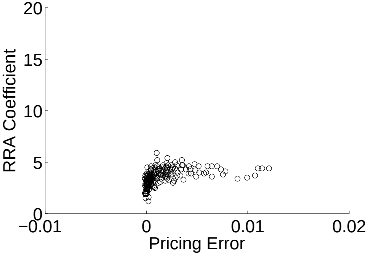
PL exponent α	3.24
Risk-free rate	1.14%
Equity premium	5.86%

- All values in annual frequency, but simulate as quarterly data.
- Since $\alpha = 3.24 < 4 = \gamma$, GMM estimation inconsistent.

Scatter plot of simulated MU RRA estimates and pricing errors (true data)



(a) Standard GMM



(b) Robust (age cohort) GMM

Multiple troughs

- In 54/200 simulations, the criterion has multiple troughs.
- The 10th and 90th percentiles for these second troughs are 6.7 and 19.4.
- In 12 of these runs, only the spurious trough is close to zero.

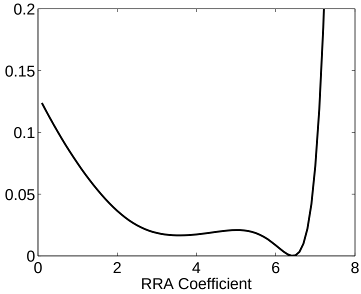
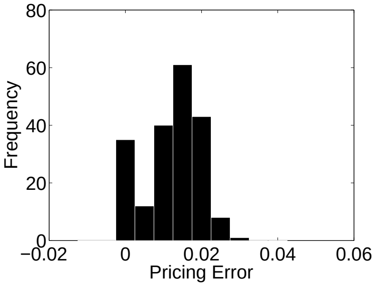
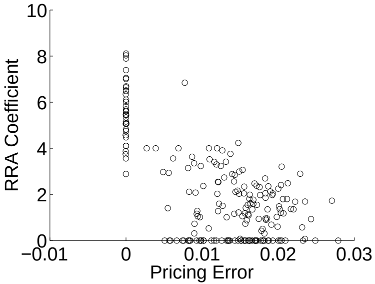


Figure: MU GMM criterion from simulation 9.

MU RRA estimates and pricing errors with false stock returns (i.i.d. copy)



(a) Histogram of pricing errors



(b) Scatter plot of simulated MU RRA estimates and pricing errors

Conclusion

- Consumption and consumption growth obey double power law with exponent ≈ 4. (We are the first to document this.)
- Power law appears to contaminate GMM estimation and mechanically zero-out pricing errors.
  $\Rightarrow$ stationary bootstrap is useful for detecting spurious results.
- With age cohort robust estimation or without outliers, we reject all asset pricing models
  $\Rightarrow$ Need better models (work in progress).
  - drop preference homogeneity.
  - include household financial frictions.

Tractable GEI with heterogeneous agents

- Agents $i = 1, \dots, I$ have identical additive CRRA preference

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

- Initial endowment (capital) $w_0 > 0$, nothing in future.
- Constant-returns-to-scale technology $j = 1, \dots, J$. Total returns to capital (productivity) for agent i between time t and $t+1$ is $A^j_{i,t+1}$. $\mathbf{A}_{i,t+1} = (A^1_{i,t+1}, \dots, A^J_{i,t+1})$ i.i.d. over time.
- Each technology have aggregate and idiosyncratic component: $A^j_{i,t+1} = a^j_{i,t+1} A^j_{t+1}$. Idiosyncratic shock $\mathbf{a}_{t+1}$ i.i.d. across agents conditional on aggregate shock $\mathbf{A}_{t+1}$.

Solving for GEI

- Let θ^j be fraction of wealth invested in technology j and

$$R_{i,t+1}(\theta) = \sum_{j=1}^{J} A^j_{i,t+1}\theta^j$$

  total returns of wealth portfolio.
- Optimal consumption and portfolio rule

$$\theta^* = \arg\max_{\theta} \frac{1}{1-\gamma} \mathrm{E}[R(\theta)^{1-\gamma}],$$
$$c(w) = \left(1 - (\beta\, \mathrm{E}[R(\theta^*)^{1-\gamma}])^{\frac{1}{\gamma}}\right) w.$$

- Asset prices $P_t = \frac{\mathrm{E}[R(\theta^*)^{-\gamma}(P_{t+1}+D_{t+1})]}{\mathrm{E}[R(\theta^*)^{1-\gamma}]}$.

Cross-sectional consumption distribution

- If agents infinitely lived, by CLT cross-sectional consumption distribution approximately lognormal.
- If agents die with probability δ between periods and capital distributed across newborn agents, by Limit Theorem ▸ ? cross-sectional consumption distribution double Pareto.
- If capital distributed according to lognormal, then cross-sectional consumption distribution dPlN ▸ ?.
- Power law exponent $\alpha = \sqrt{2\delta}/\sigma$, where σ: idiosyncratic volatility of portfolio return.

Theoretical emergence of double power law

Theorem (Toda 2012b)

Suppose that

1. $\{X_n\}_{n=1}^{\infty}$ *is a sequence of zero mean random variables such that the central limit theorem holds,*
2. $\{a_n\}_{n=1}^{\infty}$ *is a sequence such that* $N^{-1}\sum_{n=1}^{N} a_n \to a$, *and*
3. ν_p *is a geometric random variable with mean* $1/p$ *independent from* $\{X_n\}_{n=1}^{\infty}$.

Then as $p \to 0$ *the geometric sum* $p^{\frac{1}{2}}\sum_{n=1}^{\nu_p}(X_n + p^{\frac{1}{2}} a_n)$ *converges in distribution to a Laplace distribution.*

Theoretical emergence of double power law

- In the model, log wealth for an agent is the sum of period log returns over his lifespan, which is a geometric random variable.
- Using the theorem, one can show that as the time step goes to 0 log wealth converges in distribution to a Laplace distribution.
- Laplace distribution is log of double Pareto.
- See Toda (2012b) "Incomplete Market Dynamics and Cross-Sectional Distributions" for further details.

Stationary bootstrap

Explore role of power law and calculate test statistics using the stationary bootstrap of Politis & Romano (1994). $t = 1, 2, \dots, T$: time index of original sample. Construct bootstrap sample as follows:

1. Randomly pick τ_1 from $\{1, \dots, T\}$.
2. Having picked τ_n, with probability $1 - p$ set $\tau_{n+1} = \tau_n + 1$ modulo T, and with probability p pick τ_{n+1} randomly from $\{1, \dots, T\}$.
3. Here $p = 1/M(T)$, where $M(T) \to \infty$ and $M(T)/T \to 0$ as $T \to \infty$. (We set $M(T) = \sqrt{T}$.)

Account for cross-sectional uncertainty, unlike previous papers.