A Note on the Size Distribution of Consumption: More Double Pareto than Lognormal

Toda, Alexis Akira

University of California, San Diego

29 October 2015

A Note on the Size Distribution of Consumption: More Double Pareto than Lognormal

Alexis Akira Toda*

October 29, 2015

Abstract

The cross-sectional distribution of consumption is commonly approximated by the lognormal distribution. This note shows that consumption is better described by the *double Pareto-lognormal distribution* (dPlN), which has a lognormal body with two Pareto tails and arises as the stationary distribution in recently proposed dynamic general equilibrium models. dPlN outperforms other parametric distributions and is often not rejected by goodness-of-fit tests. The analytical tractability and parsimony of dPlN may be convenient for various economic applications.

Keywords: Gibrat's law, multiplicative idiosyncratic risk, inequality, power law.

JEL codes: D31, E21, G12.

1 Introduction

The cross-sectional consumption distribution is commonly approximated by the lognormal distribution, which is convenient for various economic applications. For instance, a partial list in the working paper version of Battistin et al. (2009)[1] reads: (i) lognormality implies that within cohorts, any measure of inequality, such as the Gini coefficient or the Lorenz curve, can be expressed as a function of a single scalar (the variance of log consumption), (ii) it simplifies the handling of possible measurement errors (Banks et al., 1997), and (iii) since Gabaix (1999) shows that the power law[2] for city populations may arise from an application of Gibrat's law of proportionate growth to individual cities, analogous regularities may arise in consumption from Gibrat's law. Another implication (not listed in the above paper) is that (iv) assuming lognormality of consumption enables us to derive parsimonious analytic expressions for Euler equation aggregation in heterogeneous-agent models (Constantinides and Duffie, 1996; Balduzzi and Yao, 2007). The lognormality in consumption seems to hold within age cohorts (Battistin et al., 2009).

*Department of Economics, University of California San Diego. 9500 Gilman Drive, La Jolla, CA 92093-0508, USA. Email: atoda@ucsd.edu. I thank Kieran Walsh for collaboration and discussions, and Thomas Winberry, the associate editor, and two anonymous referees for comments and feedback that significantly improved the paper.

[1] `http://fmwww.bc.edu/EC-P/wp671.pdf`

[2] Usually, "power law" refers to the property that the fraction of units with size x or larger is proportional to the power function $x^{-\alpha}$ when x is large enough. The parameter $\alpha > 0$ is called the power law exponent. The power law with $\alpha = 1$ is called Zipf's law.

In this note I take a close look at the cross-sectional consumption distribution and examine the third point above: since the cross-sectional distribution of consumption is approximately lognormal within age cohorts (which is consistent with Gibrat's law), does it obey the power law in the entire cross-section? This question itself has already been answered: Toda and Walsh (2015) show that the entire cross-sectional distribution of consumption in U.S. exhibits a Pareto tail with exponent around 4. Therefore I go one step further by incorporating the recent development in the theoretical literature on inequality. Benhabib et al. (2014) and Toda (2014) theoretically show that, in a certain class of incomplete market dynamic general equilibrium models, the cross-sectional distribution of wealth (and consumption) is *double Pareto* (Reed, 2001), which has two power law tails around its mode. Toda (2014) argues further that if agents are also heterogeneous when they are born (and the heterogeneity is characterized by a lognormal distribution), then the entire cross-sectional distribution should be *double Pareto-lognormal* (dPlN), the product of independent double Pareto and lognormal variables (Reed, 2003).

Why is it of interest to test this particular parametric distribution? There are a few reasons why testing dPlN for consumption is relevant. First, dPlN is as analytically tractable as the lognormal. A property that holds for the lognormal can often be generalized for dPlN. Second, Benhabib et al. (2014) and Toda (2014) have theoretically shown that the size distribution of consumption should be approximately lognormal within age cohorts and double Pareto or dPlN in the entire cross-section. Therefore if we find that consumption is dPlN, we have an indirect evidence suggesting that the consumption dynamics obeys the generative mechanism of dPlN. Third, recent techniques for solving heterogeneous-agent models parameterize the cross-sectional distribution to reduce computational complexity (Algan et al., 2008; Reiter, 2009; Winberry, 2015). If dPlN is a good description of reality, it may be useful in computational works because it is flexible, analytically tractable, and has only 3 or 4 parameters.

The main contribution of this note is that I document that dPlN fits remarkably well to the entire cross-sectional distribution of consumption in U.S. dPlN is often not rejected by goodness-of-fit tests, and the lognormal is almost always rejected against dPlN. Therefore consumption is heavy-tailed, and I provide an indirect evidence supporting the recent theoretical models on inequality. Another minor contribution is methodological. Although dPlN has been shown to empirically fit well the size distribution of cities (Reed, 2002; Giesen et al., 2010) and income (Toda, 2012), to the best of my knowledge dPlN has never been formally tested by goodness-of-fit tests in actual data or compared to other parametric distributions except the lognormal. Implementing goodness-of-fit tests requires a lot of coding and computing, and as a consequence, they are rarely applied in practice. Therefore my results may benefit the profession by lowering the hurdle for implementation and disciplining the empirical analysis.[3] Given that dPlN is well-grounded by theory and analytically tractable, I suggest that dPlN should be one of the benchmarks for fitting size distributions.

[3]See the Online Appendix at `https://sites.google.com/site/aatoda111` for details. All codes and data can be downloaded from the same website.

2 A simple generative mechanism of dPlN

In this section, to make the note self-contained, I present a minimal model[4] that generates the double Pareto-lognormal distribution and discuss its implications.

2.1 Gibrat's law and constant birth/death yield dPlN

Consider an economy populated by a continuum of agents indexed by $i \in I = [0, 1]$. Time is continuous, and let Y_{it} be the income of agent i at time t. Assume that income obeys Gibrat's law of proportionate growth. Letting $X_{it} = \log Y_{it}$ be the log income, Gibrat's law (in its simplest form) implies that X_{it} is a Brownian motion with drift:

$$\mathrm{d}X_{it} = g\mathrm{d}t + v\mathrm{d}B_{it},$$

where g is the drift, $v > 0$ is the volatility, and B_{it} is a standard Brownian motion which is i.i.d. across agents. Assuming that all agents start from $X_{i0} = 0$, the cross-sectional log income distribution at time t is normal: $X_{it} \sim N(gt, v^2 t)$. Hence the income distribution is lognormal, which is thin-tailed.

To get a fat-tailed distribution, suppose that agents (dynasties) "die" at a constant Poisson rate $\delta > 0$ and are reborn with initial log income $x = \log y = 0$. In steady state, the size of the cohort with age a is $\mathrm{e}^{-\delta a}$ (the density function is exponential, $\delta \mathrm{e}^{-\delta a}$), and the log income distribution for this cohort is $N(ga, v^2 a)$. Therefore the density of the entire cross-sectional distribution is the normal mixture

$$\int_0^\infty \frac{1}{\sqrt{2\pi v^2 a}} \mathrm{e}^{-\frac{(x-ga)^2}{2v^2 a}} \delta \mathrm{e}^{-\delta a} \mathrm{d}a.$$

This integral may appear complicated, but actually the integrand has an explicit primitive function (see Toda (2014) for details). The result is

$$f_{\mathrm{L}}(x) = \begin{cases} \frac{\alpha\beta}{\alpha+\beta} \mathrm{e}^{-\alpha|x|}, & (x \geq 0) \\ \frac{\alpha\beta}{\alpha+\beta} \mathrm{e}^{-\beta|x|}, & (x < 0) \end{cases} \tag{1}$$

where $-\alpha < 0 < \beta$ are solutions to the quadratic equation $\frac{v^2}{2}\zeta^2 - g\zeta - \delta = 0$. The density (1) is known as (asymmetric) *Laplace* (Kotz et al., 2001). Although the Laplace distribution was derived using a Brownian motion, Section 5 of Toda (2014) shows that it robustly arises as the limit distribution for a large class of stochastic processes. Applying the change of variable $x = \log y$ to (1), the density of the income distribution becomes

$$f_{\mathrm{dP}}(y) = \begin{cases} \frac{\alpha\beta}{\alpha+\beta} y^{-\alpha-1}, & (y \geq 1) \\ \frac{\alpha\beta}{\alpha+\beta} y^{\beta-1}, & (0 \leq y < 1) \end{cases} \tag{2}$$

which is known as *double Pareto* (Reed, 2001).[5] As its name suggests, the

[4]See Reed (2001, 2003), Benhabib et al. (2014), Toda (2014), and Toda and Walsh (2015) for similar models and Mitzenmacher (2003), Gabaix (2009), and Jones (2015) for reviews.

[5]Hence the Laplace and the double Pareto distributions have the same relation as the normal and the lognormal distributions. As an interesting historical remark, Laplace discovered the Laplace distribution in 1774, which predates by a quarter of century the discovery of the normal distribution by himself and Gauss in the early 1800s. For more historical background, see Kotz et al. (2001).

double Pareto distribution (2) has two Pareto (power law) tails, and α, β are called the power law exponents.

So far we assumed that initial log income is zero (agents are ex ante identical), but more generally we may think of situations in which agents are heterogeneous upon birth. A natural assumption is that initial log income is normally distributed, $N(\mu, \sigma^2)$. Then in steady state the log income distribution is the convolution of the normal and Laplace distributions, which is known as *normal-Laplace* (Reed and Jorgensen, 2004). The normal-Laplace distribution has four parameters, μ, σ (mean and standard deviation of the lognormal component) and α, β (power law exponents). Through the change of variable $x = \log y$, the income distribution becomes the product of independent double Pareto and lognormal variables, which is called *double Pareto-lognormal* (dPlN, Reed, 2003).

The double Pareto and lognormal distributions are special cases of dPlN by letting $\sigma \to 0$ and $\alpha, \beta \to \infty$, respectively. This is an important point because it means the lognormal distribution, which is nested within dPlN, can be tested against dPlN by the likelihood ratio test. See Reed and Jorgensen (2004) and Hajargasht and Griffiths (2013) for various estimation methods.

2.2 Economic implications of dPlN

The double Pareto-lognormal distribution has a few attractive features. First, while the density of the double Pareto has a cusp at the mode, which does not seem to show up in the cross-sectional data, dPlN is smooth.

Second, since dPlN has a lognormal body and two Pareto tails but contains only 3 (if $\alpha = \beta$) or 4 parameters (if $\alpha \neq \beta$), it is flexible yet parsimonious. This point might be convenient for numerically solving heterogeneous-agent models. In these models, the cross-sectional distribution becomes part of an aggregate state, and recent works parameterize this distribution. For example, Algan et al. (2008), Reiter (2009), and Winberry (2015) use exponential families, histograms, and normal mixtures. Since these distributions do not have fat tails but dPlN does, it might be useful in models that match micro data.

Third, because dPlN is the product of independent double Pareto and lognormal variables, which both have closed-form moments, dPlN is analytically tractable. Indeed, if Y is dPlN with parameters $(\mu, \sigma, \alpha, \beta)$, then $Y = Y_1 Y_2$ with Y_1 double Pareto (with mode 1 and power law exponents α, β) and Y_2 lognormal, so its η-th moment is (see Reed and Jorgensen (2004))

$$\mathrm{E}[Y^\eta] = \mathrm{E}[Y_1^\eta]\,\mathrm{E}[Y_2^\eta] = \begin{cases} \frac{\alpha\beta}{(\alpha-\eta)(\beta+\eta)} \mathrm{e}^{\mu\eta + \frac{1}{2}\sigma^2\eta^2}, & (-\beta < \eta < \alpha) \\ \infty. & (\text{otherwise}) \end{cases} \tag{3}$$

Since α, β describe the inequality in the top and bottom of the distribution, we might also use them in applied works on inequality (Kunieda et al., 2014; García-Peñalosa and Turnovsky, 2015) to decompose the inequality in the top and bottom.

This expression has an interesting implication for asset pricing. In the above simple model, since idiosyncratic shocks are multiplicative, the equilibrium is autarky (Constantinides and Duffie, 1996) and consumption equals income. From the Euler equation, Balduzzi and Yao (2007) show that the growth rate of the cross-sectional moment of consumption $m_{t+1} = \mathrm{E}_{t+1}[c_{i,t+1}^{-\gamma}] / \mathrm{E}_t[c_{it}^{-\gamma}]$ is a valid stochastic discount factor (SDF) up to a multiplicative constant, where $\gamma > 0$

is the relative risk aversion of agents. When consumption is lognormal, so $\log c_{it} \sim N(\mu_t, \sigma_t^2)$, then the SDF simplifies to

$$m_t^{\mathrm{BY}} = \exp\left(\Delta\left(-\gamma\mu_t + \frac{1}{2}\gamma^2\sigma_t^2\right)\right), \tag{4}$$

where $\Delta X_t = X_t - X_{t-1}$. When consumption is dPlN, using (3) for $Y = c_{it}, c_{i,t-1}$ and $\eta = -\gamma$, the SDF becomes

$$m_t^{\mathrm{dPlN}} = \exp\left(\Delta\left(-\log(1+\gamma/\alpha_t) - \log(1-\gamma/\beta_t) - \gamma\mu_t + \frac{1}{2}\gamma^2\sigma_t^2\right)\right). \tag{5}$$

Clearly (4) is a special case of (5) by letting $\alpha_t, \beta_t \to \infty$. In each case the SDF is a simple expression of the parameters. One caveat is that we need $\gamma < \beta_t$ for moments of dPlN to exist.

3 Testing dPlN for consumption

In this section I estimate and test dPlN using U.S. consumption data. I use the same data as the real, seasonally adjusted, quarterly household consumption data from the Consumption Expenditure Survey (CEX) used in Toda and Walsh (2015). The data consists of 410,788 observations from December 1979 to February 2004. Since households are surveyed every three months but in different months, on average we have about 1,400 households in each month (or 4,200 in each quarter).

3.1 Statistical model and parameter estimation

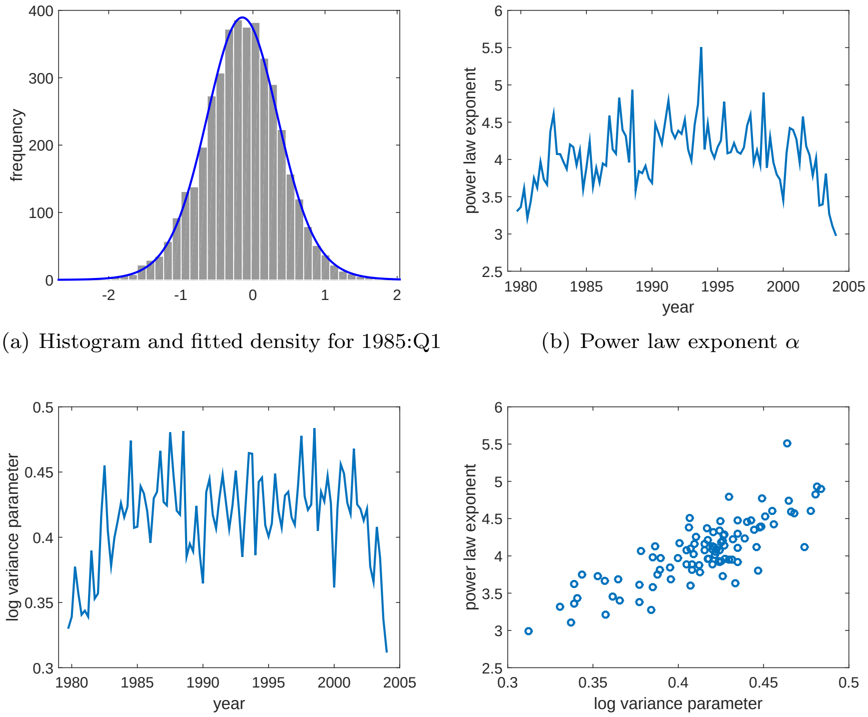
In the simple model of Section 2, we found that the cross-sectional distribution of consumption is dPlN. If log consumption has a time trend but the shocks have constant volatility, then the location parameter μ will be time-dependent but σ, α, β will be constant. It is well known that the consumption data in the CEX is subject to substantial measurement error, but dPlN survives if (the multiplicative) measurement error is lognormal because dPlN already has a lognormal component.

Now suppose that observed consumption is dPlN with parameters $(\mu_t, \sigma, \alpha, \beta)$ (or observed log consumption is normal-Laplace with the same parameters). Normalizing log consumption by subtracting the population mean, normalized log consumption is normal-Laplace with parameters $(\mu, \sigma, \alpha, \beta)$ (where μ is such that the mean is zero, so $\mu = \frac{1}{\beta} - \frac{1}{\alpha}$), which do not depend on the sample. Since the CEX samples the same households once in a quarter, by the above reasoning *monthly* data of normalized consumption has the same distribution as the *quarterly* data of normalized consumption obtained by pooling three consecutive monthly data, which contains no overlapping households. Making the sample size approximately three times larger in this way, I am taking a conservative position since it is easier to reject a particular parametric model with more data.[6]

[6]Of course, one may object pooling different data sets (although here it is theoretically justified). As a robustness check, I also perform all subsequent analysis with the original monthly data in the Online Appendix but the results are similar.

Having constructed quarterly normalized log consumption data, I estimate the normal-Laplace parameters $(\mu, \sigma, \alpha, \beta)$ for each quarter by maximum likelihood. Since the two power law exponents α, β are almost the same, I estimate the parameters of the symmetric normal-Laplace distribution by maximum likelihood. The likelihood ratio test failed to reject symmetry $(\alpha = \beta)$ in 77 out of 98 quarters at significance level 0.05. (The Online Appendix verifies through simulations that the likelihood ratio test has high power and hence the non-rejection of symmetry is not due to low power.) Therefore I choose the symmetric normal-Laplace distribution as the benchmark model for normalized log consumption.

Figure 1(a) shows the histogram of log consumption for 1985:Q1, together with the fitted symmetric normal-Laplace density plotted in the range between the minimum and the maximum log consumption (other quarters look similar). We can see that the fit is quite good. Figure 1(b) shows the maximum likelihood estimate of the power law exponent α for each quarter. The power law exponent is around 4 (the average across all quarters was 4.06) and in the range $[3, 5]$ except 1993:Q4, where $\widehat{\alpha} = 5.5$. The sample average of the standard error across all quarters was 0.31. These results are similar to Toda and Walsh (2015).



(a) Histogram and fitted density for 1985:Q1

(b) Power law exponent α

(c) Log variance parameter σ

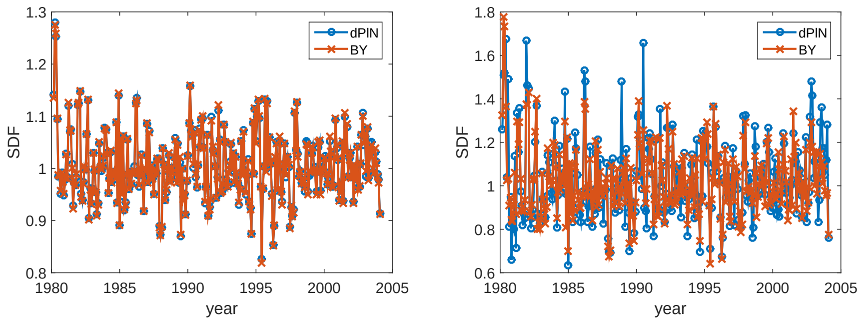
(d) Scatter plot of σ, α

Figure 1: Maximum likelihood estimation of symmetric normal-Laplace distribution to quarterly U.S. normalized log consumption data.

What is the meaning of a power law exponent of 4? For income, it is well-known that the exponent is about 2 or 3 (Toda, 2012). Therefore the simple model in Section 2—which predicts the same exponent for income and consumption—is clearly false. Thus an interesting question might be to look at the dynamics of household income and consumption and investigate how we should modify the model in order to explain the tail heaviness of income and consumption observed in the data, but is left for future research.

Figure 1(c) shows the estimated log variance parameter σ for each quarter, which shows a similar pattern to Figure 1(b). Indeed, the scatter plot of $\widehat{\sigma}, \widehat{\alpha}$ in Figure 1(d) shows a positive relation (correlation 0.81). However, this is probably an artifact of sampling error. For dPlN, tails are fatter when σ is large or α is small. Hence for a fixed sample, there is a trade-off in the fit between increasing σ and decreasing α. At the best fit parameter values, when one is large, the other also tends to be large. Indeed, I estimated the parameters from 100 simulated datasets with sample size $N = 4,000$ and dPlN parameters $(\mu, \sigma, \alpha, \beta) = (0, 0.4, 4, 4)$, and obtained similar graphs and correlation 0.89.

How economically important is it to model consumption as dPlN instead of lognormal? Figure 2 shows the Balduzzi-Yao (BY) stochastic discount factor in (4) and the dPlN SDF in (5) for relative risk aversion $\gamma = 1$ and 2.5. With low risk aversion ($\gamma = 1$, Figure 2(a)), the two SDFs are virtually identical (correlation 0.998). However, even slightly increasing the risk aversion ($\gamma = 2.5$, Figure 2(b)) makes the two SDFs quite different (correlation 0.85): the dPlN SDF is more volatile.

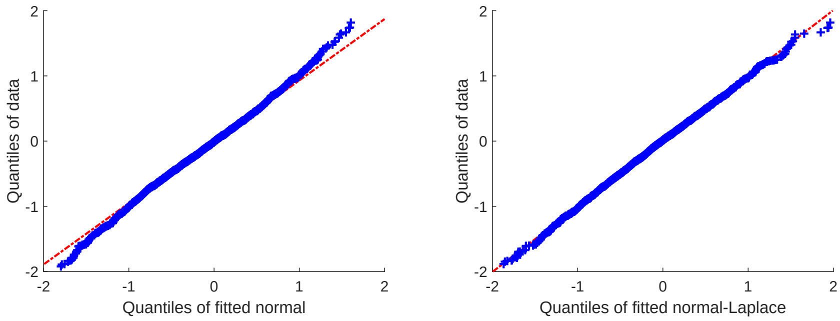


(a) $\gamma = 1$ (b) $\gamma = 2.5$

Figure 2: dPlN and BY stochastic discount factor.

3.2 Goodness-of-fit

According to Figure 1(a), the symmetric normal-Laplace distribution seems to fit the log consumption data well. We can also see that there are large positive and negative values that would be unlikely if the distribution were normal. To see this visually, Figure 3 shows the quantile-quantile plot (QQ plot) of log consumption against the fitted normal (Figure 3(a)) and normal-Laplace (Figure 3(b)). In Figure 3(a), the actual quantiles (vertical axis) are more extreme than those of the normal distribution (horizontal axis), suggesting that the log consumption distribution has heavier tails than the normal distribution. On the other hand, the QQ plot against the normal-Laplace distribution shows a virtually straight line, suggesting the good fit.



(a) Normal

(b) Normal-Laplace

Figure 3: Quantile-quantile plot of normalized log consumption. 1985:Q1.

Although a picture is worth a thousand words, to evaluate the goodness-of-fit of the double Pareto-lognormal distribution (with $\alpha = \beta$) and the lognormal distribution more formally, I perform both the Kolmogorov-Smirnov (KS) test (Massey, 1951) and the Anderson-Darling (AD) test (Anderson and Darling, 1952) and compute the P value by bootstrapping 500 times for each quarter as explained in the Online Appendix. Letting $F(x)$ be the theoretical distribution and $F_N(x)$ the empirical cumulative distribution function, the KS and AD test statistics are based on the sup and L^2 norms,

$$\sup_x |F_N(x) - F(x)| \quad \text{and} \quad \int_{-\infty}^{\infty} \frac{(F_N(x) - F(x))^2}{F(x)(1 - F(x))} \mathrm{d}F(x).$$

Note that the KS test has a low power for detecting deviations from the theoretical distribution in the tails because $F_N(x) - F(x)$ tends to zero as $x \to \pm\infty$. On the other hand, the AD test can detect deviations in the tails because the weighting function $[F(x)(1 - F(x))]^{-1}$ tends to infinity as $x \to \pm\infty$. Hence, with the Anderson-Darling test the deviations in the tails are more penalized. (The Online Appendix verifies through simulations that the KS test has indeed a low power.)

Since I am interested in the tail heaviness, for my purpose clearly the AD test is more appropriate. However, I also perform the KS test because it is widely used.[7] The double Pareto-lognormal distribution (with $\alpha = \beta$) is not rejected by the KS test at significance level 0.05 in 79 quarters out of 98. On the other hand, the lognormal distribution is rejected in 73 quarters. When using the AD test, which is more relevant because I am interested in the tail behavior, dPlN is not rejected in 64 quarters out of 98, whereas lognormal is rejected in 92 quarters. The fact that the lognormal distribution is rejected more often by the AD test suggests that the lognormal distribution fails to fit the tails of the

[7] In testing the lognormal distribution within age cohorts, Battistin et al. (2009) perform the KS test as well as tests based on skewness measures that use quantiles (but not the AD test). Since their skewness measures use only octiles (12.5 percentile, 25 percentile, etc.), their test cannot address the fit in the tails (1 percentile or 99 percentile, say). In fact, as I show below in testing lognormality against dPlN, within age cohorts the KS and AD tests reject the lognormal distribution about 30% and 50% of the time, which suggests that the tails indeed matter for evaluating the model fit.

data, which is consistent with the QQ plot in Figure 3(a). The Online Appendix shows the actual P value of these tests.

3.3 Comparison to other parametric distributions

I also compare the performance of dPlN to other parametric distributions with different tail heaviness using the Bayesian Information Criterion (BIC). (Using AIC gives similar results.) The parametric distributions that I consider are the lognormal, gamma, and generalized beta II (GB2) of McDonald (1984).[8] GB2 has four parameters a, b, p, q with density

$$f_{\mathrm{GB2}}(x) = \frac{ax^{ap-1}}{b^{ap}B(p,q)(1+(x/b)^a)^{p+q}},$$

where $b > 0$ is a scale parameter, $a, p, q > 0$ are shape parameters, and $B(p, q)$ denotes the beta function. The attractive feature of the generalized beta II distribution is that it nests a wide range of parametric distributions such as the exponential, gamma, lognormal, Weibull, and other distributions as special or limiting cases.[9] Out of 98 quarters, dPlN performed best in 78 quarters, GB2 in 15 quarters, lognormal in 5 quarters, and gamma in none. Therefore, among a large class of parametric distributions, dPlN provides the best fit to the consumption distribution.

3.4 Testing dPlN against lognormal distribution

Since the lognormal distribution is nested within the double Pareto-lognormal distribution (by letting the power law exponents α, β to infinity), we can test the lognormal distribution against dPlN by the likelihood ratio test. The test rejects the lognormal distribution at significance level 0.05 in every quarter except 1993:Q4, with P value 0.08 (which gives the largest power law exponent 5.5 in Figure 1(b)). Therefore the consumption distribution is better described by dPlN than lognormal when we look at the entire sample.

This finding does not necessarily contradict to those of Battistin et al. (2009) because they look at the consumption distribution *within age cohorts*, not the entire cross-section. Since the double power law emerges from the constant probability of the "birth/death" process as in the model in Section 2, 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, I perform the likelihood ratio test and goodness-of-fit tests for the lognormal distribution for each age cohort. The groups are household head age 30 or less, 31 to 40, 41 to 50, 51 to 60, and 60 or more. The likelihood ratio test fails to reject lognormality in 46, 38, 37, 56, and 32 quarters out of 98 for each age group, respectively. The Kolmogorov-Smirnov test fails to reject lognormality in 67, 69, 63, 72, and 43 quarters for each age group, and the Anderson-Darling test fails to reject in 51, 51, 50, 64, and 23 quarters, respectively. Therefore the lognormal distribution fits reasonably well to the cross-sectional distribution of

---

[8] I do not compare to the classic Lévy alpha-stable distribution (Mandelbrot, 1961) because (i) it does not have a closed-form density, and (ii) its power law exponent is necessarily below 2, whereas for consumption the exponent is well above 2.

[9] See McDonald (1984) and Toda (2012) for more discussion. GB2 does not nest the double Pareto or dPlN, but it does obey the double power law with exponents $\alpha = aq$ and $\beta = ap$.

consumption for each age group, in agreement with Battistin et al. (2009). My finding that the double power law emerges only in the entire cross-section and not within age cohorts is consistent with the theoretical model of Benhabib et al. (2014) and Toda (2014), in which the age distribution is geometric due to the constant probability of birth/death.

References

Yann Algan, Olivier Allais, and Wouter J. Den Haan. Solving heterogeneous-agent models with parameterized cross-sectional distributions. *Journal of Economic Dynamics and Control*, 32(3):875–908, March 2008. doi:10.1016/j.jedc.2007.03.007.

Theodore W. Anderson and Donald A. Darling. Asymptotic theory of certain 'goodness of fit' criteria based on stochastic processes. *Annals of Mathematical Statistics*, 23(2):193–212, 1952. doi:10.1214/aoms/1177729437.

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.

James Banks, Richard Blundell, and Arthur Lewbel. Quadratic Engel curves and consumer demand. *Review of Economics and Statistics*, 79(4):527–539, November 1997. doi:10.1162/003465397557015.

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.

Jess Benhabib, Alberto Bisin, and Shenghao Zhu. The distribution of wealth in the Blanchard-Yaari model. *Macroeconomic Dynamics*, 2014. doi:10.1017/S1365100514000066.

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

Xavier Gabaix. Zipf's law for cities: An explanation. *Quarterly Journal of Economics*, 114(3):739–767, 1999. doi:10.1162/003355399556133.

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

Cecilia García-Peñalosa and Stephen J. Turnovsky. Income inequality, mobility, and the accumulation of capital. *Macroeconomic Dynamics*, 19(6):1332–1357, September 2015. doi:10.1017/S1365100513000850.

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.

Gholamreza Hajargasht and William E. Griffiths. Pareto-lognormal distributions: Inequality, poverty, and estimation from grouped income data. *Economic Modelling*, 33:593–604, July 2013. doi:10.1016/j.econmod.2013.04.046.

Charles I. Jones. Pareto and Piketty: The macroeconomics of top income and wealth inequality. *Journal of Economic Perspectives*, 29(1):29–46, Winter 2015. doi:10.1257/jep.29.1.29.

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

Takuma Kunieda, Keisuke Okada, and Akihisa Shibata. Finance and inequality: How does globalization change their relationship? *Macroeconomic Dynamics*, 18(5):1091–1128, July 2014. doi:10.1017/S1365100512000843.

Benoît Mandelbrot. Stable Paretian random functions and the multiplicative variation of income. *Econometrica*, 29(4):517–543, October 1961. doi:10.2307/1911802.

Frank J. Massey, Jr. The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American Statistical Association*, 46(253):68–78, 1951. doi:10.1080/01621459.1951.10500769.

James B. McDonald. Some generalized functions for the size distribution of income. *Econometrica*, 52(3):647–663, May 1984. doi:10.2307/1913469.

Michael Mitzenmacher. A brief history of generative models for power law and lognormal distributions. *Internet Mathematics*, 1(2):226–252, 2003. doi:10.1080/15427951.2004.10129088.

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. On the rank-size distribution for human settlements. *Journal of Regional Science*, 42(1):1–17, February 2002. doi:10.1111/1467-9787.00247.

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.

Michael Reiter. Solving heterogeneous-agent models by projection and perturbation. *Journal of Economic Dynamics and Control*, 33(3):649–665, March 2009. doi:10.1016/j.jedc.2008.08.010.

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.

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 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.

Thomas Winberry. Lumpy investment, business cycles, and stimulus policy. Job Market Paper, 2015.