Economics
The Open-Access, Open-Assessment E-Journal

Discussion Paper
No. 2017-67 | September 26, 2017 | http://www.economics-ejournal.org/economics/discussionpapers/2017-67

Rawls' fairness, income distribution and alarming level of Gini coefficient

Yong Tao, Xiangjun Wu, and Changshuai Li

Abstract
The arguments that the alarming level of Gini coefficient is 0.4 are widely reported. However, to the authors´ knowledge, it is not based upon any rigid economic theories. In this paper, they show that Rawls' fairness is compatible with the standard model of competitive markets. This finding reveals that the exponential income distribution not only satisfies Pareto optimality (or efficiency) but also obeys social fairness in Rawls' sense. Therefore, the authors specify the maximal value of Gini coefficients when income follows exponential distribution as a minimal basic reference point of the alarming level (calculated as 0.5), above which efficiency and Rawls' fairness cannot be guaranteed simultaneously. Their empirical investigations show that during peaceful times, worldwide Gini coefficients approximately obey asymptotic normal distribution with a mean around 0.4, contradicting the implication of alarming level; while the two-sigma rule shows that in our sample the alarming levels are all larger than 0.5, conforming to our prediction.

JEL D31 D51 D63
Keywords Rawls' fairness; Competitive equilibrium; Income distribution; Gini coefficient

Authors
Yong Tao, College of Economics and Management, Southwest University, China
Xiangjun Wu, ✉ School of Economics, Hangzhou Dianzi University, China, wuxiangjun@hust.edu.cn
Changshuai Li, John and Walk Department of Economics, Clemson University, USA

Citation Yong Tao, Xiangjun Wu, and Changshuai Li (2017). Rawls' fairness, income distribution and alarming level of Gini coefficient. Economics Discussion Papers, No 2017-67, Kiel Institute for the World Economy. http://www.economics-ejournal.org/economics/discussionpapers/2017-67

Received September 19, 2017 **Accepted as Economics Discussion Paper** September 20, 2017
Published September 26, 2017

1 Introduction

The Gini coefficient is commonly used as a measure of income inequality among a nation's residents. Since income inequality is often regarded as the cause of social instability, the Gini coefficient is naturally identified as an important early-warning signal. For a long time, the international institutions, such as the World Bank, UN, and the news media, etc., accept that the alarming level of Gini coefficient should be set at 0.4 (see the UN report (2013)). Some authors also implicitly hint that 0.4 is a critical value for the more developed countries (MDCs) (see, e.g., Niancotti 2006; Chin and Culotta 2014). This implies that if the Gini coefficient of one country exceeds 0.4, it may confront the risk of overall social instability. However, to our knowledge, economists even cannot make clear where this alarming level comes from. In fact, it is neither derived based on any available economic theory nor confirmed by convincing empirical investigations. If 0.4 is incorrect, reducing income inequality inadequately may hinder economic development. In this paper, we will investigate whether the alarming level of 0.4 is supported by existing data sets.

Since the seminal work of Kuznets (1955), there has emerged a huge and well-developed literature on the theoretical and empirical studies of income inequality (see de la Croix and Doepke 2003; Yang 2013). The empirical support for the existence of income inequality and its importance in generating social, political and economic instabilities has promoted introducing such a variable into the models of macroeconomics. Given the broad interests in studying the effects of income inequality on macroeconomic aggregates and its determinants, economists have been constructing and testing all kinds of theories and empirical studies (see, for example, Aghion 2002; Golosov et al. 2013; Kelly 2000; Krusell et al. 2000; Kuznets 1955). However, the problems are far from settled. Increasing interest in income inequality has intensified efforts to provide the tools to measure and analyze the distribution of income. Most research is empirically related to the measurement and evaluation of the changes in income inequality. The popular approach is to use the share of total income accruing to some parts of the top income holders, such as the top 10% group, to measure the income concentration (see Alvaredo et al. 2003; Piketty 2003; Piketty and Saez 2003, 2014). Some other papers are more prone to employ the indexes, such as Gini coefficient, to gauge the overall inequality (see Hvistendahl 2014; Ravallion 2014; Xie and Zhou 2014). Anand and Segal (2008) and Chin and Culotta (2014) propose excellent discussions of the methodological issues on the measurement of inequality. A comparison between these two strands of measures, see Alvaredo (2011) and Cowell and Flachaire (2007). Unfortunately, all of these efforts do not discern how income inequality may result in social instability. For instance, what types of income distributions will guarantee social stability?

Reconciling efficiency and equity is a significant theme in the world today. Intuitively, the alarming level of Gini coefficient should be at least set at a value, above which efficiency and equity cannot be satisfied simultaneously. Then, a question arises immediately: is there a trade-off between efficiency and equity? If we identify equity with equal allocation, the answer may be "no". However, when one considers Rawls' alternative notion of equity (Ralws 1999), the so-called equal opportunity, and the answer will become affirmative. Recently, Tao (2016a) showed that reconciling "Rawls' equal opportunity" and "competitive efficiency" was feasible. It is well known that Arrow-Debreu's general equilibrium model (hereafter ADGEM) can be regarded as the standard tool for dealing with the efficient resource allocation among social members (Mas-Collel et al. 1995). The solution to ADGEM is called the competitive equilibrium, corresponding to a Pareto optimal income allocation. Here, "Pareto optimality" implies "efficiency". However, regarding the case of long-run competition, Tao (2016a) argues that the ADGEM will have multiple solutions so that no one, by Arrow's Impossibility Theorem, can clarify which solution is best. This is just the well-known "Dilemma of Social Choice" (Arrow 1963). To eliminate this dilemma, Tao shows that if one imposes Rawls' principle of fair equality of opportunity on the long-run ADGEM, an exponential income distribution, which is a set of Pareto optimal income allocations, would occur spontaneously (namely, with the largest probability) (Tao 2015, 2016a). This work reminds us that one can seek a possible alarming level of Gini coefficient in the sense of Rawls' fairness (rather than equalitarianism). Our main idea is that because the exponential distribution is a result reconciling efficiency and Rawls' fairness, we may specify the maximal value of Gini coefficient of such a distribution as a minimal basic reference point of the alarming level. Later, we will test whether such an alarming level is empirically supported by available data.

2 Model

The standard model of free markets capturing reasonable private property rights and judicial justice is called the ADGEM (Mas-Collel et al. 1995). Similar to Newtonian equations in physical world, the optimal behaviors of social members in an ideally economic world will be governed by ADGEM, which are the cornerstone of neoclassical economics (Tao 2015). The solution to ADGEM is the famous competitive equilibrium, in which income allocation is Pareto optimal. Unfortunately, if the ADGEM have multiple solutions, by Arrow's Impossibility Theorem, no one can seek the best allocation (Arrow 1963). To avoid this dilemma, Tao (2016a) proposes the paradigm of natural selection (rather than social choice), that is, "survival of the likeliest", to materialize such an evolutionism. To actualize the paradigm, Tao imposes Rawls' principle of fair equality of opportunity on ADGEM; therefore, each Pareto optimal income allocation would occur with an equal probability[1]. Then, Tao (2015, 2016a) shows that an exponential income distribution, as a set of Pareto optimal income allocations, will occur spontaneously (namely, with the largest probability) as below:

$$f_B(\varepsilon) = \begin{cases} (1/\theta)e^{-(\varepsilon-\mu)/\theta}, & \varepsilon \geq \mu \\ 0, & \varepsilon < \mu \end{cases}, \tag{1}$$

where ε denotes income, a continuous variable, μ denotes the marginal labor-capital return, and θ denotes the marginal technology return (Tao 2016a).

The exponential income distribution (1), known as "Spontaneous Economic Order" (Tao 2016a), is expressed in terms of the framework of ADGEM. In such a framework, each firm is sufficiently competitive and hence looks like a self-employed household or a small trader; therefore, the exponential income distribution (1) actually describes the income level among

[1] To see how Rawls' principle of fair equality of opportunity leads to the equal probability of equilibrium outcome, let us concentrate on Rawls' pure procedural justice (Rawls, 1999; Page 74) which aims to design the social system (or economic institutions) so that the outcome is just whatever it happens to be, at least so long as it is within a certain range. With this idea, a just economy can be regarded as a fair procedure that will translate its fairness to the (equilibrium) outcomes (Tao 2016a); thus, every social member would have no desire to oppose or prefer a certain outcome. Technically, to ensure that the economy is one of pure procedural justices, Rawls suggested considering the principle of fair equality of opportunity (Rawls, 1999; Page 76). In accordance with this principle, a fair economy implies that each outcome should be selected with equal opportunities; in other words, each outcome will then occur with an equal probability (Tao 2016a). Since ADGEM is an ideally just procedure, the fair equality of opportunity indicates that each Pareto optimal income allocation should occur with an equal probability. It is worth mentioning that here we investigate income allocation rather than wealth allocation; therefore, our result has nothing to do with merit or parentage.

households (Tao et al. 2016). Thus, we can make the theory testable using micro datasets (household income). It is worth mentioning that, due to Rawls' fairness, the exponential distribution (1) arises because the society is assumed to be ideally fair. Unfortunately, human society can never be in such an ideal state, so the exponential distribution (1) only suits for a part of the population. Yakovenko et al. (2009) employ the income data from U.S. in 1983-2000 to confirm that the income of the majority of population (lower class) obeys exponential distribution (or Boltzmann-Gibbs distribution), see Figure 1. Nirei and Souma (2007) find the same result by using income data from Japan, see Figure 2. Tao et al. (2016) study more than 60 countries all over the world, and find the exponential income distribution are highly robust. Other similar results, see Cho (2014); Jagielski and Kutner (2013); Shaikh et al. (2014). Here we emphasize that although there are many other distribution formulas to fit the households' income data, see Kleiber and Kotz (2003). These formulas lack the rigid foundation of neoclassical economics. For example, the fitting parameters of these formulas have no certain economic meanings. Therefore, no one can guarantee that each of these fitting results makes sense. In fact, by Weierstrass' Theorem in mathematical analysis, one can always employ the polynomial function to fit any strange and eccentric income data! However, compared to other distribution formulas, the exponential distribution (1) is based on ADGEM and hence owns the rigid foundation of neoclassical economics. In particular, the fitting parameters μ and θ imply certain economic meanings and have been confirmed by the OECD data (Tao 2016b).

Moreover, from Figure 1 and 2, we further notice that the income of a small fraction of population (upper class) obeys Pareto distribution (or power distribution). It is well known that Pareto distribution can be derived using some unfair rule, e.g., the rule "The rich get richer" (namely, Matthew effect) (Barabasi and Albert 1999). By the same method, Tao (2015) gets a Pareto distribution:

$$f_P(\varepsilon) = \begin{cases} \gamma a^\gamma \varepsilon^{-\gamma-1}, \varepsilon \geq a \\ 0, \qquad \varepsilon < a \end{cases}, \tag{2}$$

where ε denotes income, a continuous variable. We denote by a the income level of the entrants, $(1/\gamma)$ denotes the income share grabbed by incumbents when he is the entrants (Tao 2015).

Since the exponential distribution (1) reconciles Pareto's optimality and Rawls' fairness, we may identify it with a signal indicating social stability. Due to this, we further consider the maximal value of Gini coefficient of exponential distribution (1) as a minimal basic reference point of alarming level. By the techniques proposed in Appendix A, the Gini coefficients of the exponential distribution $f_B(\varepsilon)$ and the Pareto distribution $f_P(\varepsilon)$ are calculated as follows:

$$G_B = 1/[2(1+\mu/\theta)], \tag{3}$$

$$G_P = 1/(2\gamma - 1), \tag{4}$$

where the sign "B" denotes exponential distribution and "P" denotes Pareto distribution.

From Tao (2010, 2015, 2016a), we know that $\mu \geq 0$, $\theta \geq 0$ and $\gamma \geq 1$, so the intervals of G_B and G_P are:

$$0 \leq G_B \leq 0.5, \tag{5}$$

$$0 \leq G_P \leq 1. \tag{6}$$

As we have observed, when the income follows the exponential distribution, the world is free of the extreme inequality since G_B lies in the interval of 0 and 0.5. In contrast, Pareto distribution (2) cannot rule out the extreme inequality. Pareto distribution may ensure efficiency; but one cannot guarantee that it is fair at least in Rawls' sense. Formally, we specify the maximal value of G_B when income follows exponential distribution, which equals 0.5, as a minimal basic reference point of the alarming level, which clearly violates the international standard of 0.4. We will give the empirical evidence in the next section.

3 Data description and methodology

The main objects of our analysis are Gini coefficients, and the data employed by us comes from a sub-sample of World Bank's PovcalNet database (World Bank 2015), and the datasets are based on household microdata of various sorts. We consider cross sectional data from four years, that is: 1990, 1995, 2000 and 2005, the total observations are separately 130, 137, 139 and 140, which are representatives of the overall countries in the world. Our test approach is on the basis of the following proposition:

Proposition 1: If there are no political or economic interventions among countries, and if total number of countries is sufficiently large, Gini coefficients among countries will follow an asymptotic normal distribution.

Proof. We provide a heuristic proof; see Appendix B for details. □

Proposition 1 implies that Gini coefficients follow a stable normal distribution when sample size is large enough and adverse shocks are rare. In fact, only when the samples follow stable distributions can we make credible statistical inferences, while the conclusions based on unstable distributions are not reliable. Although our proof is heuristic, we will later see that the empirical investigations indeed support the validity of the Proposition 1.

To test our theory, we mainly resort to Jarque-Bera Chi-square statistic to examine the normality of the data, and then employ the statistical decision theory to detect the alarming level. The Jarque-Bera statistic is (Jarque and Bera 1980, 1987):

$$JB = \frac{n}{6}\left(S^2 + \frac{1}{4}(K-3)^2\right), \tag{7}$$

where n is the number of observations (or degrees of freedom in general); S is the sample skewness, and K is the sample kurtosis. If the data comes from a normal distribution, the JB statistic asymptotically has a chi-squared distribution with two degrees of freedom. The null hypothesis is a joint test of skewness and excess kurtosis being zero. As the definition shows, any deviation from zero increases the JB statistic. For small samples, the chi-squared approximation is overly sensitive, often rejecting the null when it is in fact true. Furthermore, the distribution of P-values departs from a uniform distribution and becomes a right-skewed uni-modal distribution, especially for small P-values. As a result, there is a high rate of Type I error. So, to circumvent the small sample problem, we desert the data sets where observations are small, and choose some representative years when sample size is relatively large. In these years, our samples cover all the countries the World Bank's PovcalNet collected.

To further confirm the proposition 1, some other normality tests, such as the Lilliefors test, Cramer–von Mises criterion, Watson test, Anderson–Darling test, as well as the probability plot, will also be examined for the robustness check. The results are reported in Table 2-5 and Figure 7.

Regarding the alarming level of Gini coefficient, we follow the approach of statistical decision theory, that is, the so-called three-sigma rule (Bartoszynski and Niewiadomska-Bugaj 2008). According to this rule, we are allowed to disregard the possibility of a random variable deviating from its mean by more than three standard deviations. Namely, if $X \sim N(\mu, \sigma^2)$, where μ is the mean, σ^2 is the variance, then

$$P\{|X-\mu| > 3\sigma\} = P\left\{\left|\frac{X-\mu}{\sigma}\right| > 3\right\} = 0.003. \quad (8)$$

From equation (8), we know that for any random variable sampled from a normal distribution, the probability that it deviates from its mean by more than three standard deviations is about 3‰, which is, at most, very small and negligible. This rule is helpful to drop some outliers in the sample, while at the same time without losing much information, but here our purpose is not selecting observations, so we make a compromise on the rule and use a variant–the two-sigma rule–to test our theory. In fact, the choosing of two-sigma rule is not based on our subjective willingness, but on empirical facts. King and Zeng (2001a, 2001b) find that the occurrence of wars only consists of 0.34% in international relations from a very large dataset. We treat the random variables deviating from their means more than two standard deviations as the small probability events, meaning that the occurrences of these events are barely possible (rather than impossible). Correcting equation (8), the probability of a random variable deviating from its mean by more than two standard deviations is 0.046, and in the right tail, it is 0.023. Any events that happen at a probability of either of this or lower are regarded as small probability events, and then the critical values of Gini coefficients can be calculated due to the following assumption:

Assumption 1: In times of overall peace, the instability around the world is a small probability event.

Assumption 1 states that in times of peace, only very small number of countries may experience social instability. History records show that, when the world at peace, overall economic uncertainty and large-scale turbulence are highly impossible, in statistical jargon, they are small probability events. We do not rule out the possibility of social instability, but just propose that it is not a systemic event. Equivalently, if the sample size is large, observations undergoing political instability are only negligible outliers.

By two-sigma rule, during peaceful periods, the probability that instability occurs lies around:

$$P\{|X-\mu| > 2\sigma\} = P\left\{\left|\frac{X-\mu}{\sigma}\right| > 2\right\} = 0.046.$$

For the normal distribution, we define the alarming level of Gini coefficients as $X_c = \mu + 2\sigma$. If the Gini coefficient is larger than X_c, small probability event occurs, and by Assumption 1 the country starts to show signs of instability.

4 Results

In this section, we report the empirical results in two separate parts, due to the fact that the normal distribution is rejected in the year of 1990. The results of the Jarque-Bera tests are reported in Figure 3–6, and the other normality tests are shown in Table 2–5, while the probability plots are reported in Figure 7.

4.1 World in 1990

In 1990, the normal distribution is rejected both from the Jarque-Bera test and all the other tests (see Figure 3 and Table 2). The probability plot also shows that normal distribution may not hold; see Panel A in Figure 7 for details. This result is very interesting and insightful, though, not contradicted with our theory (see Proposition 1). The history of 1990 across the world was in fact very dark and full of uncertainty, politics and economy went into chaos, a series of astonishing incidents occurred, such as the reunification of Germany, the Gulf War and the Baltic states declaring independence from the Soviet Union, et al., to name a few. When all these events reflected in the data, the result is that Gini coefficients no longer follow a stable normal distribution. Due to this, we have no way to calculate the alarming level. Although the mean of Gini coefficients is smaller than that of other years, the standard deviation is much larger, a signal of instability.

4.2 World in 1995, 2000, and 2005

When it comes to the years of 1995, 2000, and 2005, it is clear that normal distribution of the null hypothesis cannot be rejected at the 1% marginal significance level; See Figure 4-6 and Panel B to D in Figure 7, as well as the Table 3-5 for details. In fact, the P-values of all the normality tests are larger than 1%, a strong signal showing that the null hypothesis cannot be rejected. Unlike 1990, the world in these years was in the states of tranquility. More importantly, by Figure 4–6 we note that 0.4 is just, without exception, the mean of the sample. In other words, 0.4 corresponds to the most probable event. If 0.4 is set as an alarming level, during these years one must encounter that the majority of the countries in the world confront the risk of overall social instability; however, history records did not show such instability. To avoid this dilemma, the 0.4 standard must be rejected.

By the two-sigma rule, the critical values are calculated as 0.590, 0.573, and 0.560, see Table 1, they are clearly compatible with our theoretical value 0.5. From Figure 3–7, one can see that in normal years, the distributions are very stable, while in turbulent year, stable distribution breaks. Contrasting to the international standard of the alarming level, no one lies around 0.4; in fact, they are all much larger, showing that the traditional standard is intrinsically fault. In our sample, 0.4 is just the mean, corresponding to the most probable value, rather than the critical one. Thus, the alarming level of 0.4 cannot fully reflect the facts of social instability.

Although the choice of data (spanned by 5 years) makes it difficult to establish a link between economy and political uncertainty, it is possible to link above-stated coefficients and economic fluctuations, especially when comparing the social states in 1990 with the other three years. From the comparisons, the results show that our proposition of the alarming level of Gini coefficient (0.5) is credible. In summary, the empirics show that the alarming levels of Gini coefficients are all at least equal to or larger than 0.5 in our samples. So, in the practical policy-making process, when Gini coefficient is rising, but has not yet crossed the alarming level, the priorities for governments are still enhancing economic development, rather than alleviating income inequality.

5 Discussion

Due to an absence of rigid theory on traditional argument of the alarming level of Gini coefficient, we adopt Tao’s exponential income distribution which captures “competitive efficiency” and “Rawls’ fairness” simultaneously to circumvent this deficiency. To be specific, the exponential distribution not only satisfies the conditions of Pareto optimality, but also conforms to Rawls’ principle of fair equality of opportunity; therefore, it is the direct result of fair competition in a society. However, when it comes to reality, a problem arises: human society can never be ideally fair in Rawls’ sense. This implies that in countries with mature and sound legal systems, as well as democratic regimes, for the majority of people, income follows exponential distribution. While at the other end, for the minorities, income follows a non-fair distribution, according to the "the rich get richer" rule, it is Pareto distribution. These inferences have been supported by empirical investigations. The existing studies of the U.S., Japan and other countries show that the income of the majorities follows exponential distribution, while the minorities follows Pareto distribution, implying that income distribution in a society consists of two distinct parts. In particular, exponential distribution reconciles efficiency and Rawls' fairness, so we may specify the maximal value of Gini coefficient when income follows exponential distribution, which equals 0.5, as a minimal basic reference point of the alarming level.

Based on our model, we test the theory using data of Gini coefficients collected from all kinds of countries in four separate years. We first show that Gini coefficients are normally distributed in states of tranquility, while in the turbulent year, a stable normal distribution under the significance level of 5% no longer exist. This result presents an implication for seeking potential alarming level of Gini coefficient when regional or political conflicts are small probability events in a peaceful world, but when the world undergoes radical changes or turbulences, the absence of stable distribution makes the calculation of alarming level impossible. Next we calculate the alarming levels from three years' data, that is, 1995, 2000, and 2005, using statistical decision theory. The results suggest that the alarming levels are all larger than 0.5, supporting the proposal posed in our theoretical model. An interesting exception is in 1990, the Gini coefficients no longer follow normal distribution at the 5% significance level. Though not appropriate, we still informally calculated the value of the alarming level under normal distribution using the two sigma rule, to be 0.606, which implies that when a country's Gini coefficient is larger than 0.6, the society may be prone to be unstable, just like the year of 1990. Although our model fits the reality very well, a caveat also needs to be applied to the empirics: The alarming level we proposed only suits for the free market system, which ensures the free competition and equal opportunity in a large part.

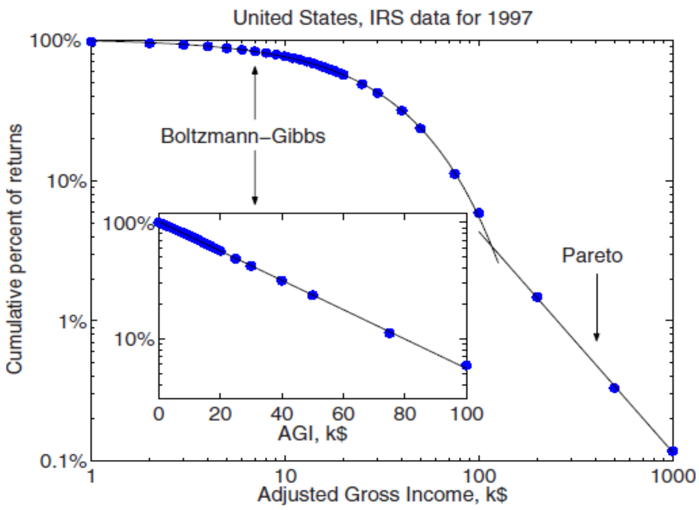

Figure 1 Reprinted from Yakovenko and Rosser (2009). Points represent the Internal Revenue Service data, and solid lines are fits to Boltzmann-Gibbs and Pareto distributions.

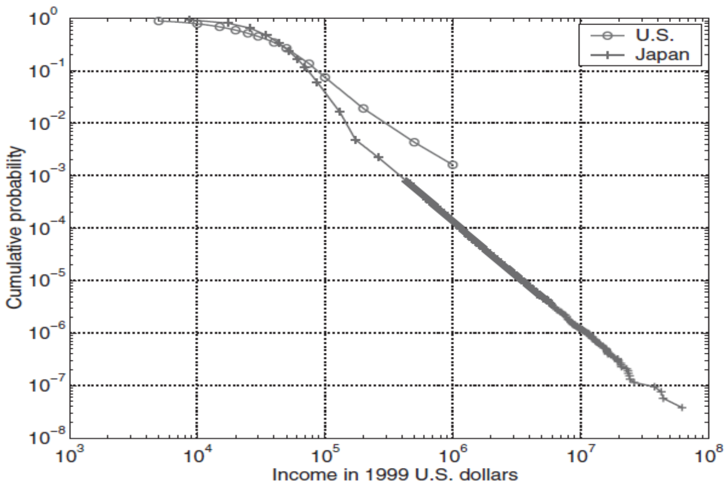

Figure 2 Reprinted from Nirei and Souma (2007). Income distributions in the U.S. and Japan in 1999.

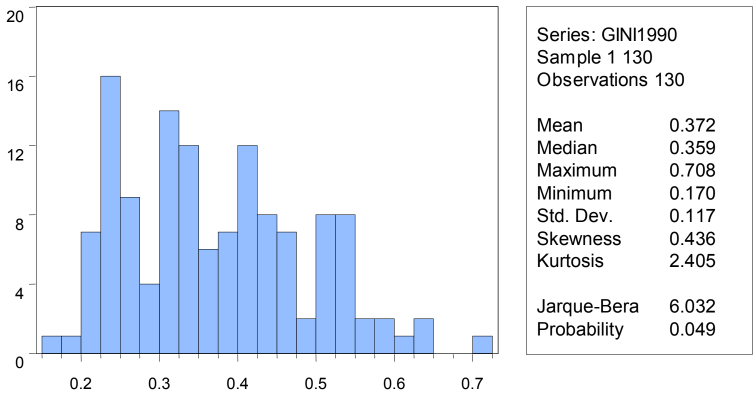

Figure 3 The histogram and summary statistics of Gini coefficient in 1990.

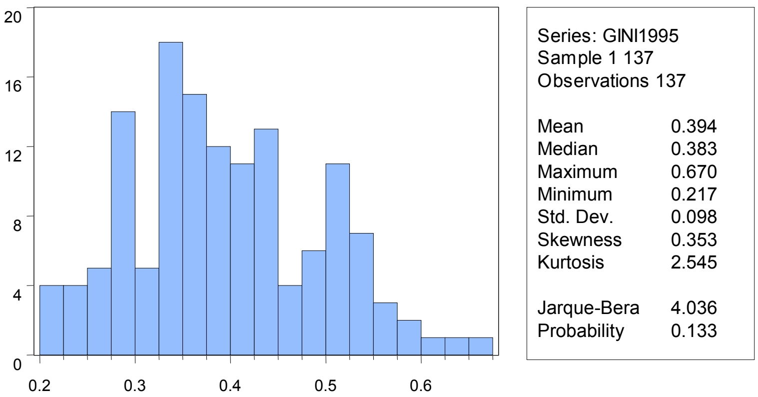

Figure 4 The histogram and summary statistics of Gini coefficient in 1995.

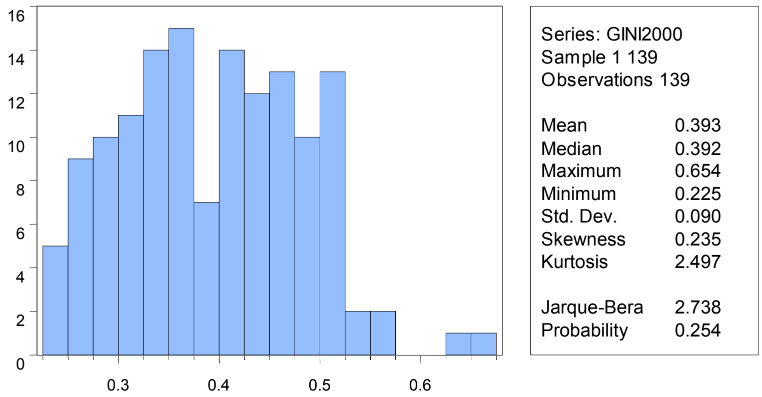

Figure 5 The histogram and summary statistics of Gini coefficient in 2000.

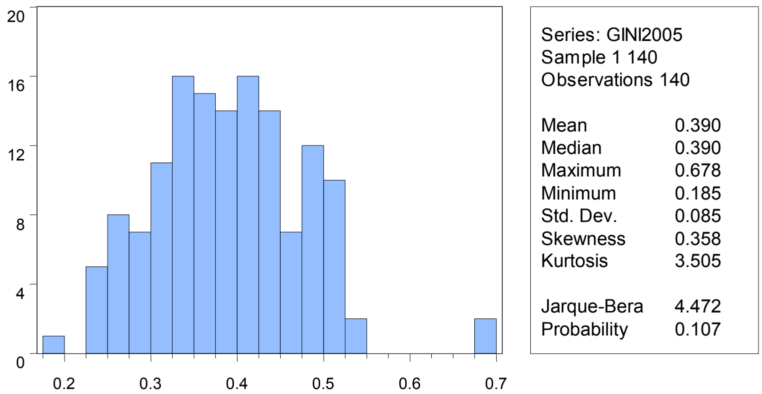

Figure 6 The histogram and summary statistics of Gini coefficient in 2005.

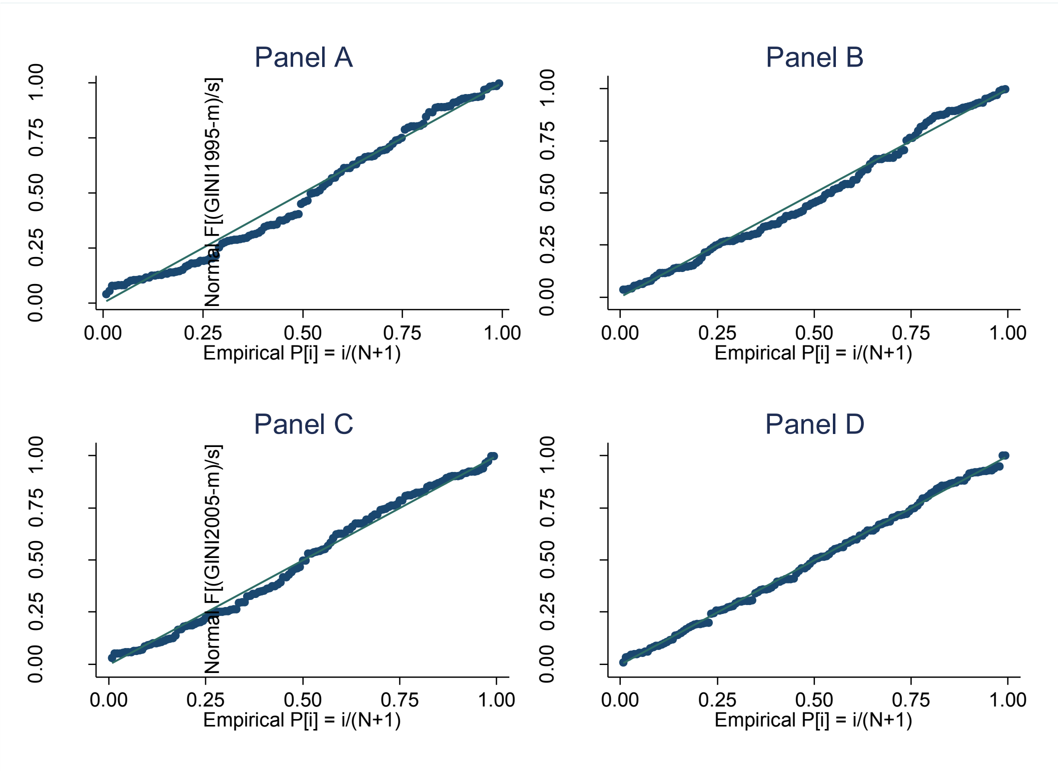

Figure 7 The probability plots of the Gini coefficients in 1990, 1995, 2000, and 2005.

Table 1 The alarming level of Gini coefficients

Year	Alarming level of Gini coefficient
1995	0.590
2000	0.573
2005	0.560

Note: The data of Gini coefficient in 1990 no longer follows normal distribution at the 5% significance level, so the alarming level is not reported.

Table 2 Normality tests in 1990

Method	Value	Adjusted Value	P value
Lilliefors Test	0.088	NA	0.015
Cramer-von Mises Test	0.196	0.196	0.006
Watson Test	0.174	0.175	0.007
Anderson-Darling Test	1.318	1.326	0.002

Table 3 Normality tests in 1995

Method	Value	Adjusted Value	P value
Lilliefors Test	0.061	NA	> 0.1
Cramer-von Mises Test	0.127	0.127	0.048
Watson Test	0.112	0.112	0.059
Anderson-Darling Test	0.777	0.782	0.043

Table 4 Normality tests in 2000

Method	Value	Adjusted Value	P value
Lilliefors Test	0.067	NA	> 0.1
Cramer-von Mises Test	0.112	0.112	0.076
Watson Test	0.109	0.109	0.065
Anderson-Darling Test	0.786	0.790	0.041

Table 5 Normality tests in 2005

Method	Value	Adjusted Value	P value
Lilliefors Test	0.037	NA	> 0.1
Cramer-von Mises Test	0.025	0.025	0.913
Watson Test	0.021	0.021	0.946
Anderson-Darling Test	0.328	0.330	0.515

Note: The “Value” reports the asymptotic test statistics, the “Adjusted Value” reports test statistics that have a finite sample correction or adjusted for parameter uncertainty (in case the parameters are estimated). The forth column reports the P-Value for the adjusted statistics.

Appendix. Proofs

Appendix A

In this appendix, we will derive the expression of Gini coefficient under exponential and Pareto distribution, respectively.

We assume that the income level x in an arbitrary country is a continuous variable, and lies in a closed interval $[a, b]$, where $a \geq 0$ and $b < +\infty$. The probability density function (PDF) is $f(x)$, with cumulative distribution function (CDF) being $F(x)$. Thus, $F(x)$ is the percentage of population whose income less than x:

$$F(x) = \int_a^x f(x)\,dt. \tag{A1}$$

Then the mean of the income $\overline{x}$ is:

$$\overline{x} = \int_a^b x f(x)\,dx. \tag{A2}$$

Under the settings assumed above, the formula for calculating the much-used Gini coefficient G can be written as (see, for example, Lambert (1993)):

$$G = \frac{2}{\overline{x}} \int_a^b x \left[F(x) - \frac{1}{2}\right] f(x) dx. \tag{A3}$$

Now we use the formula (A3) to calculate the Gini coefficients of exponential distribution and power distribution, respectively.

Substituting (1) into (A3) and by order $b \to \infty$ we obtain:

$$G_B = 1/[2(1 + \mu/\theta)]. \tag{A4}$$

Substituting (2) into (A3) and by order $b \to \infty$ we obtain:

$$G_P = 1/(2\gamma - 1). \tag{A5}$$

Appendix B

In this appendix, we provide a *heuristic* proof for the Proposition 1 as used by Gauss (1809), that is, the measurement error obeys normal distribution. Even so, we still remind the readers that our method cannot be regarded as a rigid treatment. We leave the strict proof to readers who may be interested in Proposition 1.

In a market-oriented economy, it is naturally supposed that by the "invisible hand" the economy will produce a desirable Gini coefficient. We can denote the desirable Gini coefficient by G_D. However, there are lots of different factors which force the actual Gini coefficient to deviate from the desirable one. For simplicity, here we focus mainly on that there are different resource endowments among countries so that the distribution of Gini coefficients among countries may be non-uniform. Thus, we denote by G_1, G_2,…, G_n the samples of n different countries' Gini coefficients. Our following proof is due to two assumptions.

Proof. First, we assume that G_i is independent of G_j for any $i \neq j$. This means that there are no political or economic interventions among countries.

Denote by $\varepsilon_i = G_i - G_D$ for $i = 1, \ldots, n$ the deviation of every single country's Gini coefficient from the desirable Gini coefficient. Suppose ε_i's probability density function is $f(\varepsilon_i)$, then the joint density function for all the observations is

$$L(G_D) = L(G_D;\, G_1, \ldots, G_n) = \prod_{i=1}^{n} f(\varepsilon_i) = \prod_{i=1}^{n} f(G_i - G_D) \tag{B1}$$

By the method of maximum likelihood estimation, one sets

$$\frac{d \log L(G_D)}{dG_D} = 0. \tag{B2}$$

Rearranging (B2) as

$$\sum_{i=1}^{n} \frac{f'(G_i - G_D)}{f(G_i - G_D)} = 0. \tag{B3}$$

Substituting $g(\mathrm{G}_i - G_D) = \frac{f'(G_i - G_D)}{f(G_i - G_D)}$ into (B3) yields

$$\sum_{i=1}^{n} g(G_i - G_D) = 0. \tag{B4}$$

Second, we assume that the solution governed by maximum likelihood estimation is exactly the arithmetic average, provided that $n \to \infty$. Such an assumption implies that $lim_{n \to \infty} \bar{G} = G_D$, where $\bar{G} = \frac{\sum_{i=1}^{n} G_i}{n}$. Implementing this assumption is on the basis of Aumann's famous result (Aumann 1966) which states that the competitive equilibrium always exists when the number of competitors approaches infinity. Accordingly, if we specify a competitor by a country, then we acknowledge that resource allocation between infinite countries will be Pareto optimal. Therefore, the difference between Gini coefficients is due to the difference between resource endowments. Since each country's resource endowment can only be randomly endowed by nature, we can adopt Gauss's assumption which states that the desirable value is just the average value.

So we plug the arithmetic average of Gini coefficient, $\bar{G}$, into (B4) and obtain

$$\sum_{i=1}^{n} g(G_i - \bar{G}) = 0. \tag{B5}$$

If we set $n = 2$, then we have

$$g(G_1 - \bar{G}) + g(G_2 - \bar{G}) = 0. \tag{B6}$$

It is easy to see $(G_1 - \bar{G}) = -(G_2 - \bar{G})$. Since G_1 and G_2 are arbitrary, then we get

$$g(-\varepsilon) = -g(\varepsilon) \tag{B7}$$

We set $n = m + 1$ in (B5) and meanwhile let $\varepsilon_1 = \cdots \varepsilon_m = -\varepsilon$ and $\varepsilon_{m+1} = m\varepsilon$ so that $\bar{\varepsilon} = 0$, then we have

$$\sum_{i=1}^{n} g(G_i - \bar{G}) = mg(-\varepsilon) + g(m\varepsilon). \tag{B8}$$

Substituting (B.5) and (B.7) into (B.8) yields

$$g(m\varepsilon) = mg(\varepsilon). \tag{B9}$$

The only continuous function $g(\cdot)$ satisfying (B9) is $g(\varepsilon) = c\varepsilon$, so we get

$$f(\varepsilon) = Me^{c\varepsilon^2}. \tag{B10}$$

Due to $\int_{-\infty}^{+\infty} f(\varepsilon) d\varepsilon = 1$ finally we have

$$f(\varepsilon) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\varepsilon^2}{2\sigma^2}}. \tag{B12}$$

□

It is worth mentioning that Gauss's assumption stating that the desirable value is just the average value, may not be strict. However, the stringency of Gauss's assumption has been overcome by Laplace (1820).

Acknowledges Tao acknowledges the support of the Fundamental Research Funds for the Central Universities, (Grant No. SWU1409444). Wu acknowledges the support of Natural Science Foundation of China (Grant No. 71271096) and National Excellent Doctor Fund (Grant No. 201304). We are grateful to seminar participants at Huazhong University of Science and Technology, Southwest University, Clemson University and Chongqing University for helpful comments. All errors are our own.

References

Aghion, P. (2002). Schumpeterian growth theory and the dynamics of income inequality. Econometrica, 70 (3), 855–882. http://onlinelibrary.wiley.com/doi/10.1111/1468-0262.00312/abstract

Alvaredo, F. (2011). A note on the relationship between top income shares and the Gini coefficient. Economics Letters, 110 (3), 274–277. http://www.sciencedirect.com/science/article/pii/S0165176510003460

Alvaredo, F., Atkinson, A., Piketty, T., Saez, E. (2003). The top 1 percent in international and historical perspective. Journal of Economic Perspectives, 27 (3), 3–20. https://www.aeaweb.org/articles?id=10.1257/jep.27.3.3

Anand, S., Segal, P. (2008). What do we know about global income inequality? Journal of Economic Literature, 46 (1), 57–94. https://www.aeaweb.org/articles?id=10.1257/jel.46.1.57

Arrow, K. (1963). Social Choice and Individual Values. John Wiley& Sons.

Aumann, R. (1966). Existence of competitive equilibria in markets with a continuum of traders. Econometrica, 34 (4), 1-17. https://www.econometricsociety.org/publications/econometrica/1966/01/01/existence-competitive-equilibria-markets-continuum-traders

Barabasi, A., Albert, R. (1999). Emergence of scaling in random networks. Science, 286, 509–512. http://science.sciencemag.org/content/286/5439/509

Bartoszynski, R., Niewiadomska-Bugaj, M. (2008). Probability and statistical inference, 2ed, John Wiley & Sons.

Biancotti, C. (2006). A polarization of inequality? The distribution of national Gini coefficients 1970–1996. Journal of Economic Inequality, 4, 1–32. https://link.springer.com/10.1007/s10888-005-4889-7

Chin, G., Culotta, E. (2014). The science of inequality: what the numbers tell us? Science, 344, 818-821. http://science.sciencemag.org/content/344/6186/818

Cho, A. (2014). Physicists say it's simple. Science, 344 (6186), 828. http://science.sciencemag.org/content/344/6186/828

Cowell, F., Flachaire, E. (2007). Income distribution and inequality measurement: The problem of extreme values. Journal of Econometrics, 141 (2), 1044–1072. http://www.sciencedirect.com/science/article/pii/S0304407607000036

de la Croix, D., Doepke, M. (2003). Inequality and growth: why differential fertility matters. American Economic Review, 93 (4), 1091–1113. https://www.aeaweb.org/articles?id=10.1257/000282803769206214

Gauss, C. (1809). Theoria Motus Corporum Celestium. Hamburg, Perthes et Besser. Translated as Theory of motion of the heavenly bodies moving about the sun in conic sections (trans. C. H. Davis), Boston, Little, Brown 1857. Reprinted: New York, Dover 1963, 1809.

Golosov, M., Maziero, P., Menzio, G. (2013). Taxation and redistribution of residual income inequality. Journal of Political Economy, 121 (6), 1160–1204.
http://www.journals.uchicago.edu/doi/10.1086/674135

Hvistendahl, M. (2014). While emerging economies boom, equality goes bust? Science, 344 (6186), 832–835.
http://science.sciencemag.org/content/344/6186/832

Jagielski, M., Kutner, R. (2013). Modelling of income distribution in the European Union with the Kokker-Planck equation. Physica A, 392 (9), 2130-2138.
http://www.sciencedirect.com/science/article/pii/S0378437113000629

Jarque, C., Bera, A. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. Economics Letters, 6 (3), 255–259.
http://www.sciencedirect.com/science/article/pii/0165176580900245

Jarque, C., Bera, A. (1987). A test for normality of observations and regression residuals. International Statistical Review, 55 (2), 163–172.
http://www.jstor.org/stable/1403192

Kelly, M. (2000). Inequality and crime. Review of Economics and Statistics, 82 (4), 530–539.
http://www.mitpressjournals.org/doi/abs/10.1162/003465300559028

King, G., Zeng, L. (2001a). Logistic regression in rare events data. Political Analysis, 9 (2), 137-163.
https://doi.org/10.1093/oxfordjournals.pan.a004868

King, G., Zeng, L. (2001b). Explaining rare events in international relations. International Organization, 55 (3), 693–715.
https://doi.org/10.1162/00208180152507597

Kleiber, C., Kotz, S. (2003). Statistical size distributions in economics and actuarial sciences, John Wiley & Sons, Inc., Hoboken, New Jersey.

Krueger, D., Perri, F. (2006). Does income inequality lead to consumption inequality? evidence and theory. Review of Economic Studies, 73, 163–193.
https://doi.org/10.1111/j.1467-937X.2006.00373.x

Krusell, P., Ohanian, L., Rios-Rull, J., Violante, G. (2000). Capital-Skill complementarity and inequality: a macroeconomic analysis. Econometrica, 68 (5), 1029–1053.
http://onlinelibrary.wiley.com/doi/10.1111/1468-0262.00150/abstract

Kuznets, S. (1955). Economic growth and income inequality. American Economic Review, 45 (1), 1–28.
http://www.jstor.org/stable/1811581

Lambert, P. (1993). The distribution and redistribution of income: A mathematical analysis, 2ed, Manchester University Press.

Laplace, P. (1820). Théorie analytique des probabilities, 3ed, Courcier, Paris.

Mas-Collel, A., Whinston, M., Green, J. (1995). Microeconomic Theory, Oxford University Press.

Nirei, M., Souma, W. (2007). A two factor model of income distribution dynamics. Review of Income and Wealth, 53 (3), 440–459.
http://onlinelibrary.wiley.com/doi/10.1111/j.1475-4991.2007.00242.x/abstract

Piketty, T. (2003). Income inequality in France, 1901–1998. Journal of Political Economy, 111 (5), 1004–1042.
http://www.journals.uchicago.edu/doi/10.1086/376955

Piketty, T., Saez, E. (2003). Income inequality in the United States, 1913-1998. Quarterly Journal of Economics, 118 (1), 1–39.
https://doi.org/10.1162/00335530360535135

Piketty, T., Saez, E. (2014). Inequality in the long run. Science, 344 (6186), 838–843.
http://science.sciencemag.org/content/344/6186/838

Ravalllion, M. (2014). Income inequality in the developing world. Science, 344 (6186), 851–855.
http://science.sciencemag.org/content/344/6186/851

Rawls, J. (1999). A Theory of Justice (Revised Edition), Harvard University Press.

Shaikh, A., Papanikolaou, N., Wienera, N. (2014). Race, gender and the econophysics of income distribution in the USA. Physica A, 415, 54-60.
http://www.sciencedirect.com/science/article/pii/S0378437114006153

Tao, Y. (2010). Competitive market for multiple firms and economic crisis. Physical Review E, 82, 1–8.
https://journals.aps.org/pre/abstract/10.1103/PhysRevE.82.036118

Tao, Y. (2015). Universal laws of human society's income distribution. Physica A, 435, 89-94.
http://www.sciencedirect.com/science/article/pii/S0378437115004161

Tao, Y. (2016a). Spontaneous economic order. Journal of Evolutionary Economics, 26, 467-500.
https://link.springer.com/article/10.1007/s00191-015-0432-6

Tao, Y. (2016b). An Index Measuring the Deviation of Real Economies from General Equilibria: Evidence from the OECD Countries. Available at SSRN.

Tao, Y., Wu, X. J., Zhou, T., Yan, W. B., Huang, Y. Y. X., Yu, H., Mondal, B. and Yakovenko, V. M. (2016). Universal Exponential Structure of Income Inequality: Evidence from 60 Countries. Journal of Economic Interaction and Coordination. R & R.

UNRISD, (2013). Inequalities and the Post-2015 Development Agenda: A Concept Note. Geneva.

World Bank, (2015). World Bank's PovcalNet database. http://data.worldbank.org/

Xie ,Y., Zhou, X. (2014). Income inequality in today's China. PNAS, 111 (19), 6928–6933.
http://www.pnas.org/content/111/19/6928.short

Yakovenko, V., Rosser, J. (2009). Statistical mechanics of money, wealth, and income. Reviews of Modern Physics, 81, 1703–1725.
https://journals.aps.org/rmp/abstract/10.1103/RevModPhys.81.1703

Young, A. (2013). Inequality, the urban-rural gap, and migration. Quarterly Journal of Economics, 128 (4), 1727–1785.
https://doi.org/10.1093/qje/qjt025