[[Draft. Incomplete]]

6×10^9: Some dynamics of global inequality and growth

by

Danny Quah *
LSE Economics Department

December 1999

* Portions of this paper are based on research that has been generously supported by the British Academy, the ESRC, the MacArthur Foundation, and the Mellon Foundation. I thank Clarissa Yeap for efficient research assistance.

6×10^9:
Some dynamics of global inequality and growth

by

Danny Quah
LSE Economics Department

December 1999

ABSTRACT

This paper combines "emerging twin peaks" distribution dynamics for percapita incomes across countries, with personal income distributions within countries over time. The result is a picture of worldwide income distribution dynamics across people. The paper finds that for determining world inequalities, the forces assuming first-order importance are those macroeconomic ones that determine cross-country patterns of growth and convergence. Inequality across individuals worldwide remains a critical issue of increasing concern. However, the relation between a country's growth performance and its within-country inequality plays only a small role in global inequality dynamics.

Keywords: convergence, distribution dynamics, Gini coefficient, income distribution, poverty

JEL Classification: C23, D31, F43

Communications to: D. Quah, LSE, Houghton Street, London WC2A 2AE.

Tel: +44/0 207 955-7535, Email: `dquah@econ.lse.ac.uk`
(URL) `http://econ.lse.ac.uk/~dquah/`

6×10^9:
Some dynamics of global inequality and growth
by
Danny Quah *
LSE Economics Department
December 1999

1 Introduction

It is a truism easily overlooked that whether inequality is positively or negatively related to economy-wide growth says nothing about whether inequality across people worldwide is rising or falling as the entire world economy continues to grow.

Similarly, since in every economy incomes are differentially distributed across people, whether countries are converging to the same Solow steady state—much less to ad hoc individually different growth paths—says nothing about whether the poor in the world are catching up with the rich.

This paper has two goals. First, it seeks to quantify global inequality dynamics.[1] Has global inequality increased? What is the distribution of income across people worldwide? What are possibilities for the future?

A second goal is, in part, methodological: This is to quantify what information can be usefully extracted from inequality indicators, such as Gini coefficients, quantile or decile shares, and so on. At one extreme, these indicators provide exactly the information they say they

* Portions of this paper are based on research that has been generously supported by the British Academy, the ESRC, the MacArthur Foundation, and the Mellon Foundation. I thank Clarissa Yeap for efficient research assistance.

[1] It shares this goal with at least three other recent papers, Bourguignon and Morrison (1999), Milanovic (1999), and Heston and Summers (1999). All four papers have approached this same question, although using different tools and perspectives.

do, i.e., each is a summary representation of the entire income distribution. But much economic research attempts to use these statistics (or distribution functionals) not only in that narrow way, but instead as correlates across countries in, say, a growth regression.

Thus, for instance, a well-established literature studies how income inequality affects output growth, within a given economy (see, e.g., the survey in Bénabou, 1996). That work uses cross-country observations to assess validity of the theory's predictions. Each country is considered an independent, autarkic unit. National policy variation—in taxation, education systems, transfers, and so on—makes the empirical exercise a sensible one, for the former is what *defines* an economy in that research. In addressing the first goal set out above, this paper also sheds light on the quantitative relevance of such empirical analyses.

The questions I posed above on global inequality differ from those addressed in most of the now-traditional work on inequality and growth. Implicit in both, however, is a view that inequality is a public policy issue. Examining global inequality simply recognizes the universality of that social concern: Should UK inequality, say, be more of an issue than inequality in the US? Or inequality in France, Honduras, or China?

This paper considers the question by examining the dynamics of the income distribution across the six billion people on the planet. It merges information on the distribution dynamics of per capita incomes *across* countries (e.g., Jones 1997, Pritchett 1997, and Quah 1993, 1997), and that on income distributions *within* countries over time (e.g., Deininger and Squire 1996 and Li, Squire and Zou 1998).

The principal findings are as follows:

1. Given recent historical patterns, the positive effects of economic growth on individual incomes overwhelm any potential negative impacts from realistic, i.e., not historically unusual, increases in inequality.
2. If populations were kept constant, worldwide aggregate economic growth moves upwards the entire distribution of individual incomes. Those dynamics can in any one year be expected

to move more than 100 million people [**CONFIRM**] with less than US$2 a day out of that state of poverty. Within-country inequalities would have to increase at practically-inconceivable rates to overcome this tendency of aggregate economic growth to enrich everyone. In the event, between 1980 and 1992, the actual outturn in population growth and inequality dynamics led to there being [**CONFIRM**] 1 billion fewer people with incomes of less than US$2 a day.

3. Between 1980 and 1992 within-country inequalities rose in some countries, and fell in others. Overall, however, inequality across people worldwide rose—almost entirely due to changes in the distribution of per capita incomes across countries. Thus, while inequality should remain an issue of great social concern, the greatest changes in world inequality can be expected to derive from (macroeconomic) growth and convergence of country per capita incomes, not from (microeconomic) changes in within-country inequalities.

These results emphasize that for determining world inequalities, the forces assuming first-order importance are those macroeconomic ones that determine cross-country aggregate patterns of growth and convergence. Certainly, inequality across individuals worldwide remains a critical issue of potentially increasing concern. However, the relation between a country's growth performance and its within-country inequality–whether inequality promotes growth or reduces it; whether growth increases inequality or lowers it—plays only a small role in determining global inequality dynamics.

Before continuing, it might help to be explicit on some of the numbers that make up these findings. Abstract from details here—for those, see Section **3** below. Consider a simplified, two-country world.

Two economies China and India together account for over one-third of the world's population. In 1980 the Gini coefficients of income inequality in both China and India were equal at 0.32. By 1992 inequality had risen in China to a Gini coefficient of 0.38; that in India remained constant. The rise in China's inequality was certainly not monotone; indeed, inequality *fell* to a Gini value of 0.26 in 1984. Neither was inequality in India constant throughout. The Gini coefficient fluctuated over time, but simply happened to take the same values at both the beginning and end of the sample period.

Over these 12 years, per capita income in China increased at an annual rate of 3.6%; that in India, at 3.1%. Thus, inequality rose in the economy that grew faster.

Given observations on per capita incomes in addition (and under further assumptions described in Section **3** below) one can estimate that the fraction of the population with income less than US$2 per day (in 1992 dollars, purchasing power parity adjusted) fell in China from 37—54% of the population in 1980 to 14–17% in 1992, and in India from 48–62% in 1980 to 12–19% in 1992. (There is a range rather than just one number, as the estimate varies with auxiliary assumptions.) Given historical population dynamics, the world comprising just these two countries—where cross-sectionally growth is unrelated to inequality levels but positively related to changes in inequality—between 374 and 604 million people grew out of having less than US$2 a day in incomes.

For this calculation the absolute threshold level of US$2 a day matters because the process is not linear. While the entire distribution is moving upwards, different parts of the distribution do so at different rates. At the same time, however, the transition is *smooth* throughout the range of income levels. It is not that people just below a given threshold level are being moved just above it, and then left there. The entire distribution moved.

This result—about 500 million past US$2 a day—combines the effects of a number of different factors. Among those forces at work are an increasing population, economic growth in the aggregate, and changes in inequality. We can isolate some of these factors by posing a couple of different questions. First, from the perspective of 1980, with inequality and population held constant, how many people were exiting the state of having incomes less than US$2 a day? Second, how fast would inequality have had to be growing to keep the same fraction of the population below US$2 a day? For China, the answer to the first is 33 million people a year (or 3.4% of the 1980 population);

for India, 17 million (2.4%). Thus in 1980, if nothing else were to change, China and India through aggregate growth alone would have expected in a single year to move past the US$2-a-day threshold a total number of people equal to the entire UK population.

The answers to the second are 8.3% per year proportional growth in the Gini coefficient for China; 8.8% for India. (Since China grew faster in the aggregate, one might have expected the Gini to need to grow faster there too. However, since it was already richer in 1980, the multiplier from per capita to required Gini growth rates is correspondingly smaller. See equation (20) in **3.4** below.) Both these would have entailed increases in inequality from a Gini coefficient of 0.32 to 0.35 *in one year*.

While temporary changes in inequality of this magnitude exist in the data, sustaining them for over a decade is unknown. Even the UK's widely remarked-on Gini coefficient rise of 50% over 15 years took a proportional growth rate of only under 2.8% per year.[2]

What has happened to the global distribution of individual incomes, in this artificial two-country world comprising only China and India? It is well-known that Gini coefficients of inequality neither compose across subgroups nor decompose in them. But that is irrelevant for the approach taken in Section **3** below. From the worldwide income distribution in equation (1), estimated using (5) and section **3.2**, it is easy to see that inequality across people worldwide has increased. However, it is also easy to verify that almost all of that rise in global inequality has been due to the originally richer econ-

[2] These calculations, obviously, do not take into account that changes in inequality might influence aggregate growth. But then to be clear, neither do they assume that aggregate growth affects inequality, one way or another. Indeed, there is no presumption of one factor causing the other nor of parameters and coefficients remaining invariant under policy interventions. Instead, this paper attempts only to document the empirical regularities that transpired historically. Calculations like those in the previous paragraph are not "as if" policy experiments. They serve only to provide an interpretable window into a complicated distribution dynamics.

omy, China, also growing faster in aggregate. If the timepaths of per capita incomes for China and India had been switched, but maintaining the actual changes in inequality within each of the two countries, global inequality would have *fallen* instead.

How robust are these conclusions? In the calculations for them, I have used needed additional assumptions together with knowledge of inequality and incomes in just individual economies to:

- construct statements about *global* inequality;
- assess rates of population movements past absolute income levels;
- estimate growth rates of inequality required to maintain a population at incomes below a given threshold (poverty) level.

Some of these statements are compositional; others are dynamic; all require additional maintained assumptions. In other words, we wish to find out about what is happening across the world—not just in a single country; we wish to understand what is happening over time, not just at one snapshot point in time.

The rest of this paper makes explicit those assumptions I have used to build up a picture of the dynamics of global inequality and growth. It extends the calculations to include all the economies for which reliable data are available, i.e., to more than a two-country world. And it does so under a range of alternative maintained assumptions, to examine robustness.

Section **2** considers related literature, and says where the current work situates relative to others. Section **3** lays out the assumptions, estimation techniques, and analytical calculations that underly the empirical findings in Section **4** (and described in the example of the current Section). Finally, Section **5** concludes.

2 Some stylized facts and related literature

A large literature on inequality and economic growth poses the question, Is inequality positively or negatively related to growth?

Interest in this issue goes at least as far back as the Kuznets curve, suggesting the relation is first positive then negative, as the level of aggregate income rises. It will be convenient below to call relations between inequality and growth, broadly, as *Kuznets-type empirical regularities*. Even if newer, more precise facts could not have appeared, literally, in Kuznets (1955), this shorthand expression is useful and evocative.

Recent research has focused on explicit mechanisms that might explain such empirical regularities on inequality and growth. Examples include work on political economy (Alesina and Rodrik, 1994; Perotti, 1996; Persson and Tabellini, 1994) incomplete markets (Aghion and Bolton, 1997; Bénabou, 1996; Galor and Zeira, 1993) and many others. This earlier work, to put matters crudely, concerns whether the relation between inequality and growth is positive or negative in sign.

Recent empirical research (Deininger and Squire, 1996; Li, Squire and Zou, 1998) sheds a different perspective on the relation between inequality and growth. That work reorients interest away from solely the *sign* of that relation, but instead onto its *magnitude*. An important result this work is that in a sample of over 570 observations covering 49 countries, 90% of total variation in inequality is due to variation across countries; only a small percentage is due to variation in time. Put differently, as the macroeconomic development and growth process evolves for a given country, and growth rates rise and fall over time, little change occurs, on average, in that economy's cross-section income inequality.

A related finding for developing countries alone (Ravallion and Chen, 1997, p. 379) is that "higher rates of growth in average living standards are associated with higher rates of poverty reduction. The adverse distributional effect of recent growth in a number of the developing countries has not been strong enough to change the conclusion that growth has benefited the poor. For the developing countries as a whole, there is no significant trend distributional effect for or against the poor."

[[**INCOMPLETE**]]

Bring into the discussion
Heston and Summers (1999),
Milanovic (1999),
Atkinson (1998),
Snower (1998),
Galor and Tsiddon (1997),
Ravallion (1997),
Berry, Bourguignon and Morrison (1983)
Bourguignon and Morrison (1999),
Schultz (1998)
Ray (1998)
Barron and Sheu (1991)

3 Distribution dynamics across people

If data existed on individual incomes accruing to different economic agents, at each point in time, then the empirical analysis would be straightforward. One can directly estimate the entire income distribution across agents on the planet, and characterize its dynamics through time. The problem, however, is that such data are unavailable and are unlikely to be produced anytime soon.

We develop here an alternative empirical framework that is general, flexible, and convenient to use. It is designed to be capable of incorporating a wide range of alternative distributional hypotheses, and a variety of measurements on different characteristics of income inequality. Thus, the empirical analysis is intended to apply readily as more and better data on income inequality characteristics become available.

We seek to uncover characteristics of the global distribution of income across individuals. We know characteristics of income distributions *within* countries, over time for a number of countries. A traditional approach then to analyzing inequalities across progressively larger subsets of individual incomes—proceeding up from yet finer subgroups—is to ask if the inequality index *aggregates* (e.g., Milanovic, 1999). The approach I take here differs. It begins from noting that if we had the actual distribution $F_{j,t}$ for economy j at time t, where the population size is $P_{j,t}$, then the worldwide income distribution $F_{W,t}$, in a world of economies $j = 1, 2, \ldots, N$, is

$$F_{W,t}(y) = P_{W,t}^{-1} \sum_{j=1}^{N} F_{j,t}(y) \times P_{j,t}, \tag{1}$$

with the world population

$$P_{W,t} = \sum_{j=1}^{N} P_{j,t}.$$

Differentiating (1) with respect to y gives the implied density for the worldwide distribution of income as the weighted average of individual country income distribution densities:

$$f_{W,t}(y) = \sum_{j=1}^{N} f_{j,t}(y) \times (P_{j,t}/P_{W,t}) \, . \tag{2}$$

3.1 Estimating individual income distributions

Given the quantities on the right of equation (2) the worldwide income distribution is straightforward to calculate. However, the individual distributions $F_{j,t}$ are, generally, unknown. Instead, typically, we have data on a number of diverse functionals of them—e.g., Gini coefficients, quintile shares, averages, and so on. This subsection describes obtaining an estimate for F_j from data on such functionals.

Since the remainder of this section concentrates on what happens with a single economy, the j subscript is taken as understood and deleted to ease notation.

Fix an economy j. Suppose in each period t, we observe realizations on (P_t, X_t), where P is the population size and $X_t \in \mathbb{R}^d$ is a d-dimensional vector of functionals of the underlying unobservable income distribution F_t and population P_t. For example, when the first entry of X_t is the average or per capita income, then

$$X_{1,t} = \int_{-\infty}^{+\infty} y \, dF_t(y) = \int_{0}^{+\infty} y \, dF_t(y).$$

Let $(\mathbb{R}, \mathcal{R})$ denote the pair comprised of the real line $\mathbb{R}$ together with the collection $\mathcal{R}$ of its Borel sets. Let $\mathbf{B}(\mathbb{R}, \mathcal{R})$ denote the Banach space of bounded finitely-additive set functions on the measurable space $(\mathbb{R}, \mathcal{R})$ endowed with total variation norm:

$$\forall\ \varphi \text{ in } \mathbf{B}(\mathbb{R}, \mathcal{R}): \qquad |\varphi| = \sup \sum_k |\varphi(A_k)|,$$

where the supremum in this definition is taken over all

$$\{A_k : j = 1, 2, \ldots, n\}$$

finite measurable partitions of $\mathbb{R}$.

Distributions on $\mathbb{R}$ can be identified with probability measures on $(\mathbb{R}, \mathcal{R})$. Those are, in turn, just countably-additive elements in $\mathbf{B}(\mathbb{R}, \mathcal{R})$ assigning value 1 to the entire space $\mathbb{R}$. Let $\mathfrak{B}$ denote the Borel sigma-algebra generated by the open subsets (relative to total variation norm topology) of $\mathbf{B}(\mathbb{R}, \mathcal{R})$. Then $(\mathbf{B}, \mathfrak{B})$ is another measurable space.

We can write the vector of potentially-observable functionals as a collection

$$\mathbf{T}_l : (\mathbf{B} \times \mathbb{R}, \mathfrak{B} \times \mathcal{R}) \rightarrow (\mathbb{R}, \mathcal{R}),\ l = 1, 2, \ldots, d$$

(where $\mathfrak{B} \times \mathcal{R}$ denotes the sigma-algebra generated by the Cartesian product of $\mathfrak{B}$ and $\mathcal{R}$). Thus, for distribution F_t associated with probability measure $\varphi_t \in (\mathbf{B}, \mathfrak{B})$,

$$X_{l,t} = \mathbf{T}_l(\varphi_t, P_t), \quad l = 1, 2, \ldots, d. \tag{3}$$

Without loss or ambiguity, we will also write $\mathbf{T}_l(F_t, P_t)$ to denote the right hand side of (3). Write $\mathbf{T}$ to denote the vector of observed functionals, i.e.,

$$\mathbf{T}(F_t, P_t) = (\mathbf{T}_1(F_t, P_t), \mathbf{T}_2(F_t, P_t), \ldots, \mathbf{T}_d(F_t, P_t))'.$$

Assume, finally, that the distribution F_t is known up to a p-dimensional vector $\theta_t \in \mathbb{R}^p$,

$$F_t = F(\cdot|\theta_t) \stackrel{\text{def}}{=} \mathsf{F}_{\theta_t}. \tag{4}$$

(In equation (4) the symbol F is used to mean a number of different mathematical objects, but this will be without ambiguity, as the context will always be revealing.)

Equation (4) restricts in two distinct ways. First, the functional form F_t is assumed known. Second, time variation in the sequence of distributions F_t is assumed mediated entirely through the finite-dimensional parameter vector θ_t.

If for *some* θ_t^*, distribution F_{θ_t} is the true model, then

$$\mathbf{T}_l(\mathsf{F}_{\theta_t^*}, P_t) = X_{l,t},\ \ l = 1, 2, \ldots, d.$$

At fixed t, define the estimator $\hat{\theta}_t$ for θ_t^* as

$$\hat{\theta}_t \stackrel{\text{def}}{=} \mathop{\arg\min}_{\theta \in \mathbb{R}^p} \ (\mathbf{T}(\mathsf{F}_\theta, P_t) - X_t)' \Omega \left(\mathbf{T}(\mathsf{F}_\theta, P_t) - X_t\right), \quad \Omega\ d \times d \text{ positive definite.} \tag{5}$$

Each different weighting matrix Ω—including, notably, the identity matrix—produces a different estimator. Under standard regularity conditions (as in GMM or related analogue estimation, e.g., Hansen, 1982 or Manski, 1988), each Ω-associated estimator is consistent when X_t is itself replaced with a consistent estimator for the underlying population quantity. Moreover, defining the minimand

$$Q_{X_t}(\theta) = (\mathbf{T}(\mathsf{F}_\theta, P_t) - X_t)' \, \Omega \left(\mathbf{T}(\mathsf{F}_\theta, P_t) - X_t\right), \tag{6}$$

and denoting $\theta_{t,0}$ as the probability limit of (5), standard reasoning using

$$\hat{\theta}_t - \theta_{t,0} = - \left(\left. \frac{d^2 Q}{d\theta d\theta'} \right|_{\theta_{t,0}} \right)^{-1} \left. \frac{dQ}{d\theta} \right|_{\theta_{t,0}}$$

allows a limit distribution theory for these estimators, provided the quantities X_t have a characterizable distribution around their underlying population counterparts.

Using θ_t from the estimating equation (5) in (4) gives an estimator for F_t in each economy j. Plugging the result for each j in turn into (1)–(2) gives an estimator for the worldwide distribution of income. Tracking $\theta_{j,t}$ as they evolve through time then gives worldwide individual income distribution dynamics.

Section **3.4** below will provide some explicit analytically worked-out examples of this procedure. Section **4** will describe empirical results from applying the procedure to the widest extent of data available.

3.2 Alternative functionals $\mathbf{T}_l$

This subsection provides examples of some candidate functionals $\mathbf{T}_l$. When observations on them are available—as assumed in the notation of section **3.1** above—they are readily used in estimating and characterizing the distributions $F_{j,t}$. Conversely, if they are not observable but we have an estimate of $F_{j,t}$, then estimates for $\mathbf{T}_l$ can, instead, be induced.

For *mean* or *per capita income*, we take

$$\mathcal{E}(\mathsf{F}, P) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} y \, d\mathsf{F}(y). \tag{7}$$

The *Gini coefficient* is standard in analysis of income inequality. Associate with it the functional

$$\mathcal{I}_G(\mathsf{F}, P) \stackrel{\text{def}}{=} \left[2^{-1}\mathcal{E}(\mathsf{F})\right]^{-1} \int_{-\infty}^{\infty} \left(\mathsf{F}(y) - \frac{1}{2}\right) y \, d\mathsf{F}(y) \tag{8}$$

(see, e.g., Cowell, 2000).

A different set of functionals standard in inequality analyses is the set of *cumulative quintile shares*. To define these, set for integer i from 1 to 4,

$$Y_{0.2i}(\mathsf{F}, P) \stackrel{\text{def}}{=} \sup_{y \in \mathbb{R}} \{y \,|\mathsf{F}(y) \leq 0.2i\} \tag{9}$$

$$S_{0.2i}(\mathsf{F}) \stackrel{\text{def}}{=} \left(\int_{-\infty}^{Y_{0.2i}(\mathsf{F},P)} y \, d\mathsf{F}(y) \right) \times \mathcal{E}(\mathsf{F}, P)^{-1}. \tag{10}$$

The first of these, equation (9), defines the $(20 \times i)$-th percentile income level; the left-hand side is also known as the i-th quintile.

The pair (9)–(10) generalizes to arbitrary percentile shares, but in practice the more general versions are rarely used (see, however, (11), (12), and (16) below).

Concepts (8)–(10) are those traditionally used in studies on inequality. Reliable observations on them are now widely available across time and economies (Deininger and Squire, 1996).

Recently, Milanovic (1999) has used household data to construct *within-decile average incomes* across many different countries. These fit within our framework as follows. Define

$$Y_{0.1i}(\mathsf{F}, P) \stackrel{\text{def}}{=} \sup_{y \in \mathbb{R}} \{y \,|\mathsf{F}(y) \leq 0.1i\}, \quad i = 0, 1, \ldots, 9, \tag{11}$$

and let

$$\begin{aligned} \mathcal{E}_{0.1i}(\mathsf{F}, P) &\stackrel{\text{def}}{=} \int_{Y_{0.1\times(i-1)}}^{Y_{0.1i}} y \, d\mathsf{F}(y), \quad i = 1, \ldots, 9, \\ \mathcal{E}_1(\mathsf{F}, P) &\stackrel{\text{def}}{=} \int_{Y_{0.9}}^{\infty} y \, d\mathsf{F}(y). \end{aligned} \tag{12}$$

Similar to (9) above, equation (11) defines the $(10 \times i)$-th percentile income level, with the left-hand side also known as the i-th decile. The analysis in Milanovic (1999) can thus be merged with that below if we use the decile averages $\mathcal{E}_{0.1i}$ from (12) (or even the deciles themselves $Y_{0.1i}$ in (11)) as candidate $\mathbf{T}_l$'s.

Yet other ways to extract or summarize information from (F, P) are relevant when interest lies in poverty specifically (e.g., Ravallion, 1997; Ravallion and Chen, 1997; World Bank, 1990) Fix a low but otherwise arbitrary level of income $\underline{Y}$, and let:

$$HC_{\underline{Y}}(\mathsf{F}, P) \stackrel{\text{def}}{=} \mathsf{F}(\underline{Y}) = \int_{-\infty}^{\underline{Y}} d\mathsf{F}(y). \tag{13}$$

Equation (13) gives a *poverty headcount index*, i.e., the fraction of population below a given income level $\underline{Y}$. We record also the absolute size of the population with those incomes:

$$P_{\underline{Y}}(\mathsf{F}, P) \stackrel{\text{def}}{=} P \times \mathsf{F}(\underline{Y}). \tag{14}$$

Finally, define:

$$PGI_{\underline{Y}}(\mathsf{F}, P) \stackrel{\text{def}}{=} \frac{\int_{-\infty}^{\underline{Y}} y \, d\mathsf{F}(y)}{\underline{Y}}. \tag{15}$$

This is a *poverty gap index*, i.e., a (normalized) average income distance from a given income level $\underline{Y}$.

When researchers are interested in whether a gap is emerging between groups of high-income and low-income individuals, a concept more useful than just inequality is polarization (e.g., Esteban and Ray, 1994; Quah, 1993, 1997; Wolfson, 1994) To obtain a functional that captures such an effect, follow the notation of (9) and let $Y_{0.5}$ denote the *median*

$$Y_{0.5}(\mathsf{F}, P) \stackrel{\text{def}}{=} \sup_{y \in \mathbb{R}} \left\{ y \, | \mathsf{F}(y) \leq \frac{1}{2} \right\}, \tag{16}$$

and then, using (7), (8), and (16), define a *polarization index*

$$Pz(\mathsf{F}, P) \stackrel{\text{def}}{=} \left[(1 - \mathcal{I}_G)\mathcal{E} - \frac{\int_{-\infty}^{Y_{0.5}} y \, d\mathsf{F}(y)}{\int_{-\infty}^{Y_{0.5}} d\mathsf{F}(y)} \right] \times \frac{2}{Y_{0.5}}. \tag{17}$$

The first term in square brackets is the Gini-adjusted per capita income; the second is the average level of incomes below the median (this is a special case of a conditional expectation that will appear again below). The greater this separation, the higher will be the value taken by the polarization index in (17).

All the functionals so far considered—apart from $P_{\underline{Y}}$ in (14)—vary only with the distribution F, and not the size of the population P. The next functional takes both into account; it describes a dynamic property of the evolving distributions. From the headcount index (13), one might be interested in the rate of flow of people past the fixed income level $\underline{Y}$. This is

$$\begin{aligned} Fl_{\underline{Y}}(\mathsf{F}_{\theta_t}, P_t) &\stackrel{\text{def}}{=} -\frac{d}{dt} \left(\mathsf{F}_{\theta_t}(\underline{Y}) \cdot P_t \right) \\ &= -\left[P_t \frac{d}{dt} \mathsf{F}_{\theta_t}(\underline{Y}) + \mathsf{F}_{\theta_t}(\underline{Y}) \frac{dP_t}{dt} \right]. \end{aligned} \tag{18}$$

Equation (18) shows interaction among a range of factors, including in particular per capita income growth $\dot{\mathcal{E}}/\mathcal{E}$ and static, point-in-time inequality $\mathcal{I}_G$. We will use this simultaneous relationship below in sections **3.4.1** and **3.4.2**. Using different techniques, it is exactly this interaction that Ravallion (1997) studies for developing countries, using household survey data with direct observations on $Fl_{\underline{Y}}$.

The examples above should by certainly not be viewed to be exhaustive. I have given explicit $\mathbf{T}_l$ calculations only for those functionals readily found in the empirical literature and for which observations are available. As progressively more refined income-distribution data are constructed, the reasoning here is easily extended to take those into account.

3.3 Distribution F as organizing principle

As the discussion makes clear, the approach in this paper is to use the distribution dynamics in $F_{W,t}$ as the core concept around which we organize all subsequent discussion. Equation (1) is the key compositional relation from individual economies to the world. All induced statistics—Gini coefficients, poverty headcounts, poverty gap indexes, polarization indexes, and so on—derive from it. In this exercise, it is not key whether those statistics retain compositional integrity, or have an axiomatic justification, or satisfy other reasonable criteria. They are not special in this analysis. We use them below because they are easily interpretable and are standard in discussions on income distributions, thus allowing to reduce the dimensionality of (the information in) estimated distribution dynamics. As formulated here, when independently available, these statistics can be used to augment the estimation (5); when not, they can be straightforwardly derived from an estimate of $F_{j,t}$. Everything centers on the distributions.

Admittedly, backing out estimates of individual-economy distributions $F_{j,t}$—as in equation (5)—might be viewed as a contrived problem. If a researcher had the original individual-level incomes data, then $F_{j,t}$ (and thus $F_{W,t}$) could be estimated directly by standard methods (e.g., Milanovic, 1999; Silverman, 1981). One should never need to construct any of (8)–(18), and go through (5), to characterize the distribution $F_{j,t}$. It is because such individual-level data are not readily available—instead statistical agencies have earlier calculated and made available only different, aggregative statistics of the underlying data—that we are led to estimation by (5).

By the same token, one might wish to take care not to view θ as "deep structural parameters" in any sense of the term. Instead, a useful perspective is to treat the θ's as simply convenient ways—hyperparameters—to keep control on the high-dimensional calculations that would be otherwise involved in tracing through distribution dynamics. The analysis in this paper is obviously not one that sets out to test a multivariate regression or simultaneous equations model. It studies historical tendencies, not—to a large degree—the effects of artificial growth paths and inequality dynamics.

Standard econometric analysis of (5)–(6) allows consistency and limit distribution results for the hyperparameters θ. Measurement errors in the data X_t, in sample, do not logically pose any difficulties. However, whether X_t can be guaranteed to converge to underlying population quantities, and in a manner where the limiting distribution can be characterized falls outside the domain of analysis in this paper.

Finally, to state the obvious, this approach is one that makes sense when the individual distributions $F_{j,t}$ are comparable. If they are not, then the whole enterprise of trying to study worldwide inequality is flawed from the beginning, regardless of the approach taken.

3.4 Induced statistics and parametric examples

We now turn to some explicit parametric examples to provide intuition for the remainder of the analysis. In describing the distribution dynamics, it is useful to establish some additional notation.

Suppose that in a given economy per capita income $\mathcal{E}$ increases at a positive constant proportional growth rate:

$$\dot{\mathcal{E}}/\mathcal{E} = \xi > 0. \tag{19}$$

We will wish to compare dynamically evolving income distributions against a fixed (feasible and low, but otherwise arbitrary) threshold income level $\underline{Y}$. One statistic we will be concerned with in particular is the rate of flow of people past $\underline{Y}$, i.e., equation (18). We will be interested in the value of (18) when inequality, as measured by the Gini coefficient $\mathcal{I}_G$ say, is held constant. Alternatively, we will be interested in finding how fast $\mathcal{I}_G$ has to change to set (18) to zero.

Write F_θ to denote a parametrized income distribution function, and let f_θ be its associated density function:

$$\mathsf{F}_\theta(y) = \int_{-\infty}^{y} \mathsf{f}_\theta(\tilde{y})\, d\tilde{y}, \quad y \in \mathbb{R}.$$

Any given distribution also implies the conditional expectation function

$$\mathsf{E}_\theta \left(Y \,\middle|\, Y \text{ in set } \mathcal{A} \right) = \frac{\int_{\mathcal{A}} y \, d\mathsf{F}_\theta(y)}{\int_{\mathcal{A}} d\mathsf{F}_\theta(y)}.$$

This is the expectation of a random variable Y, distributed F_θ, conditional on Y falling in set $\mathcal{A}$ of possible values.

I will abuse notation by using subscripts such as $\mathsf{N}(\theta)$, $\mathsf{L}(\theta)$, or $\mathsf{P}(\theta)$ to the functions F, f, and E, to denote specific functional forms—in this case the Normal, the log Normal, and the Pareto Type 1, distributions, respectively. In the general case (with no explicit functional form restriction), the subscript will be simply θ.

To begin discussing explicitly parametrized distributions, record that the Normal distribution characterized by mean θ_1 and variance θ_2 has density

$$\mathsf{f}_{\mathsf{N}(\theta)}(y) = (2\pi\theta_2)^{-1/2} \times \exp\left\{ -\frac{1}{2\theta_2}(y-\theta_1)^2 \right\}, \quad \theta_2 > 0.$$

The *standard Normal* sets $\theta_1 = 0$ and $\theta_2 = 1$ so that then

$$\mathsf{F}_{\mathsf{N}(0,1)}(y) = \int_{-\infty}^{y} (2\pi)^{-1/2} \exp\left\{ -\frac{1}{2}\tilde{y}^2 \right\} d\tilde{y}.$$

3.4.1 Log Normal

The log Normal distribution is widely used in traditional studies of personal income distributions. Its density is

$$\mathsf{f}_{\mathsf{L}(\theta)}(y) = (2\pi\theta_2)^{-1/2} \cdot y^{-1} \times \exp\left\{-\frac{1}{2\theta_2}(\log y - \theta_1)^2\right\}, \quad \theta_2 > 0, y > 0.$$

For this distribution the **T** functionals in (7)–(10) of section **3.2** are:

$$\mathcal{E}(\mathsf{F}_{\mathsf{L}(\theta)}) = \exp(\theta_1 + \frac{1}{2}\theta_2),$$
$$\mathcal{I}_G(\mathsf{F}_{\mathsf{L}(\theta)}) = 2 \times \mathsf{F}_{\mathsf{N}(0,1)}(\,\theta_2^{1/2}/\sqrt{2}\,) - 1,$$
$$S_{0.2i}(\mathsf{F}_{\mathsf{L}(\theta)}) = \mathsf{F}_{\mathsf{L}(\theta_1+\theta_2,\theta_2)}(Y_{0.2i}(\mathsf{F}_{\mathsf{L}(\theta)})),$$

with

$$Y_{0.2i}(\mathsf{F}_{\mathsf{L}(\theta)}) = \exp\left\{\mathsf{F}^{-1}_{\mathsf{N}(0,1)}(0.2i) \cdot \theta_2^{1/2} + \theta_1.\right\}$$

An alternative expression for the cumulative quintile share is

$$S_{0.2i}(\mathsf{F}_{\mathsf{L}(\theta)}) = \mathsf{F}_{\mathsf{N}(0,1)}\left(\frac{\log Y_{0.2i} - (\theta_1 + \theta_2)}{\theta_2}\right)$$
$$= \mathsf{F}_{\mathsf{N}(0,1)}\left(\mathsf{F}^{-1}_{\mathsf{N}(0,1)}(0.2i) \cdot \theta_2^{-1/2} - 1\right).$$

If estimation (5) used only $\mathcal{E}$ and $\mathcal{I}_G$, and ignored information on other elements of **T** (or if those observations were unavailable), then an exact analytical formula for the estimator can be given:

$$\hat{\theta}_2 = \left[\mathsf{F}^{-1}_{\mathsf{N}(0,1)}\left((\mathcal{I}_G + 1)/2\right)\right]^2 \times 2,$$
$$\hat{\theta}_1 = \log \mathcal{E} - \hat{\theta}_2/2.$$

These can be used, in any case, as starting values in an iterative solution to (5).

We can then also give explicit formulas for some of the dynamics:

$$\dot{\mathcal{E}}/\mathcal{E} = \dot{\theta}_1 + \frac{1}{2}\dot{\theta}_2,$$

$$\dot{\mathcal{I}_G}/\mathcal{I}_G = \frac{\mathsf{f}_{\mathsf{N}(0,1)}([\theta_2/2]^{1/2})}{2\mathsf{F}_{\mathsf{N}(0,1}([\theta_2/2]^{1/2}) - 1} \cdot (\theta_2/2)^{1/2} \times \dot{\theta}_2/\theta_2,$$

and

$$\frac{d}{dt}\mathsf{F}_{\mathsf{L}(\theta)}(\underline{Y}) = \int_0^{\underline{Y}} \frac{d}{dt}\mathsf{f}_{\mathsf{L}(\theta)}\, dy.$$

(The Pareto case below will permit explicit calculation for all the dynamics of interest, in particular, for all the numerical results in Section **1**. Other distributional hypotheses will, as with the log Normal, require at least some of the results calculated numerically as closed-form expressions will be intractable.)

When $\mathcal{I}_G$ is held fixed, we have that $\dot{\theta}_2$ is zero. Then

$$\dot{\theta}_1 = \dot{\mathcal{E}}/\mathcal{E} = \xi,$$

so that for any fixed y,

$$\begin{aligned}\frac{d}{dt}\mathsf{f}_{\mathsf{L}(\theta)}(y) &= -(2\pi\theta_2)^{-1/2} \cdot y^{-1} \exp\left\{-\frac{1}{2\theta_2}(\log y - \theta_1)^2\right\} \\ &\quad \times (-\theta_2^{-1}) \cdot (\log y - \theta_1)(-\dot{\theta}_1) \\ &= \theta_2^{-1/2}\mathsf{f}_{\mathsf{L}(\theta)}(y) \times \left(\frac{\log y - \theta_1}{\sqrt{\theta_2}}\right)\dot{\theta}_1.\end{aligned}$$

But then,

$$\begin{aligned}-\frac{d}{dt}\mathsf{F}_{\mathsf{L}(\theta)}(\underline{Y}) &= -\theta_2^{-1/2}\left(\int_0^{\underline{Y}} \left(\frac{\log y - \theta_1}{\sqrt{\theta_2}}\right)\mathsf{f}_{\mathsf{L}(\theta)}(y)\right) \times \xi \\ &= -\theta_2^{-1/2}\mathsf{E}_{\mathsf{N}(0,1)}\left(Z \;\middle|\; Z \leq \frac{\log \underline{Y} - \theta_1}{\sqrt{\theta_2}}\right) \\ &\qquad \times \mathsf{F}_{\mathsf{N}(0,1)}\left(\frac{\log \underline{Y} - \theta_1}{\sqrt{\theta_2}}\right) \cdot \xi.\end{aligned}$$

With fixed inequality at a constant $\mathcal{I}_G$, this expression says that the flow of population past a given threshold level $\underline{Y}$ is proportional to the aggregate growth rate ξ. The constant of proportionality, moreover, is easily calculated from knowledge of θ.

The value of $\dot{\mathcal{I}}_G/\mathcal{I}_G$ that sets the flow $d\mathsf{F}_{\mathsf{L}(\theta)}(\underline{Y})/dt$ to zero can be obtained only by numerical simulation. Section **4** does this below.

3.4.2 Pareto (Type 1)

A different widely-used parametrization for personal income distributions is the Pareto (Type 1) distribution:

$$\mathsf{F}_{\mathsf{P}(\theta)}(y) = 1 - (\theta_1 y^{-1})^{\theta_2}, \quad \theta_1 > 0,\ y \geq \theta_1,\ \theta_2 > 1,$$

with density

$$\mathsf{f}_{\mathsf{P}(\theta)}(y) = \begin{cases} 0 & \text{if } y \leq 0, \\ \theta_2(\theta_1 y^{-1})^{\theta_2} y^{-1} & \text{otherwise.} \end{cases}$$

The implied **T** functionals in (7)–(10) of section **3.2** then are:

$$\begin{aligned} \mathcal{E}(\mathsf{F}_{\mathsf{P}(\theta)}) &= (\theta_2 - 1)^{-1}\theta_2\theta_1, \\ \mathcal{I}_G(\mathsf{F}_{\mathsf{P}(\theta)}) &= (2\theta_2 - 1)^{-1}, \\ Y_{0.2i}(\mathsf{F}_{\mathsf{P}(\theta)}) &= \mathsf{F}_{\mathsf{P}(\theta_1, \theta_2 - 1)}(S_{0.2i}) \end{aligned}$$

with

$$S_{0.2i}(\mathsf{F}_{\mathsf{P}(\theta)}) = (1 - 0.2i)^{-1/\theta_2} \cdot \theta_1.$$

As with the log Normal above (similarly having two parameters), an exact formula for the estimator (5) is available when only $\mathcal{E}$ and $\mathcal{I}_G$ are observed:

$$\begin{aligned} \hat{\theta}_2 &= (1 + \mathcal{I}_G{}^{-1})/2, \\ \hat{\theta}_1 &= (1 - \hat{\theta}_2^{-1})\mathcal{E}. \end{aligned}$$

In this case the dynamics in θ and ($\mathcal{E}$, $\mathcal{I}_G$) can be easily seen to be related by:

$$\begin{aligned} \dot{\mathcal{E}}/\mathcal{E} &= \frac{\dot{\theta}_1}{\theta_1} - (\theta_2 - 1)^{-1}\frac{\dot{\theta}_2}{\theta_2}, \\ \dot{\mathcal{I}}_G/\mathcal{I}_G &= \left(\frac{-2\theta_2}{2\theta_2 - 1}\right)\frac{\dot{\theta}_2}{\theta_2}. \end{aligned}$$

Moreover, direct calculation shows

$$-\frac{d}{dt}\mathsf{F}_{\mathsf{P}(\theta)}(\underline{Y}) = \frac{d}{dt}\left[\left(\frac{\theta_1}{\underline{Y}}\right)^{\theta_2}\right]$$
$$= \left(1 - \mathsf{F}_{\mathsf{P}(\theta)}(\underline{Y})\right)\theta_2$$
$$\times \left[\frac{\dot{\theta}_1}{\theta_1} + \log\left(\frac{\theta_1}{\underline{Y}}\right)\frac{\dot{\theta}_2}{\theta_2}\right].$$

When inequality in the form of $\mathcal{I}_G$ is held fixed, we have

$$\frac{\dot{\theta}_1}{\theta_1} = \frac{\dot{\mathcal{E}}}{\mathcal{E}} = \xi$$

and

$$-\frac{d}{dt}\mathsf{F}_{\mathsf{P}(\theta)}(\underline{Y}) = \left(1 - \mathsf{F}_{\mathsf{P}(\theta)}(\underline{Y})\right)\theta_2 \cdot \xi.$$

Alternatively, to fix $\mathsf{F}_{\mathsf{P}(\theta)}(\underline{Y})$ instead, we require

$$\dot{\theta}_1/\theta_1 = -\log(\theta_1/\underline{Y})\,\dot{\theta}_2/\theta_2,$$

or

$$\dot{\theta}_2/\theta_2 = -\left[\log(\theta_1/\underline{Y}) + (\theta_2 - 1)^{-1}\right]^{-1}\xi.$$

To achieve this, in turn, we need that

$$\dot{\mathcal{I}}_G/\mathcal{I}_G = \left(\frac{2\theta_2}{2\theta_2 - 1}\right)\left[\log(\theta_1/\underline{Y}) + (\theta_2 - 1)^{-1}\right]^{-1}\xi \qquad (20)$$

Equation (20) shows, at a given aggregate growth rate ξ, the rate of change in inequality required to hold fixed the proportion of the population below income $\underline{Y}$. The increase in $\mathcal{I}_G$ is proportional to ξ. When $\underline{Y}$ is sufficiently low, i.e., when

$$\mathsf{F}_{\mathsf{P}(\theta)}(\underline{Y}) < 1 - \exp\left\{\frac{-\theta_2}{\theta_2 - 1}\right\}$$

(which happens to be the situation of interest), the constant of proportionality is necessarily positive.

For the purposes of this paper, the log Normal and Pareto cases are interesting only because they permit explicit (closed-form) analyses of the distribution dynamics of interest. They provide intuition for how the general case will work. In the latter, typically only numerical solutions are available.

4 Empirical Results

Applying.

[[**INCOMPLETE**]]

5 Conclusion

This paper has provided a framework for analyzing the dynamics of global inequality and economic growth. It combined the "emerging twin peaks" description of cross-country per capita income growth and convergence with data on the evolution of personal income distributions within countries.

The result is a picture of worldwide income distribution dynamics across people. The paper finds that for determining world inequalities, the forces assuming first-order importance are those macroeconomic ones that determine cross-country patterns of growth and convergence. Certainly, inequality across individuals worldwide remains a critical issue of increasing concern. However, the relation between a country's growth performance and its within-country inequality plays only a small role in global inequality dynamics.

References

Aghion, Philippe, and Patrick Bolton (1997) “A trickle-down theory of growth and development with debt overhang,” *Review of Economic Studies* 64(2), 151–172, apr

Alesina, Alberto, and Dani Rodrik (1994) “Distributive politics and economic growth,” *Quarterly Journal of Economics* 109(2), 465–490, May

Atkinson, Anthony B. (1998) “The distribution of income in industrialized countries,” In “Income Inequality: Issues and Policy Options” Federal Reserve Bank of Kansas City pp. 11–32

Barron, Andrew R., and Chyong-Hwa Sheu (1991) “Approximation of density functions by sequences of exponential families,” *Annals of Statistics* 19(3), 1347–1369, September

Bénabou, Roland (1996) “Inequality and growth,” In *Macroeconomics Annual,* ed. Ben Bernanke and Julio Rotemberg, vol. 11 (NBER and MIT Press) pp. 11–74

Berry, Albert, François Bourguignon, and Christian Morrison (1983) “Changes in the world distribution of income between 1950 and 1977,” *Economic Journal* 93(370), 331–350, June

Bourguignon, François, and Christian Morrison (1999) “The size distribution of income among world citizens: 1820–1990,” Working Paper, DELTA, Paris, June

Cowell, Frank (2000) “Measurement of inequality,” In *Handbook of Income Distribution,* ed. Anthony B. Atkinson and Francois Bourguignon, vol. 1 (North Holland Elsevier Science) chapter 2. Forthcoming

Deininger, Klaus, and Lyn Squire (1996) “A new data set measuring income inequality,” *World Bank Economic Review* 10(3), 565–591, September

Esteban, Joan-María, and Debraj Ray (1994) "On the measurement of polarization," *Econometrica* 62(4), 819–851, July

Galor, Oded, and Daniel Tsiddon (1997) "Technological progress, mobility, and economic growth," *American Economic Review* 87(3), 363–382, June

Galor, Oded, and Joseph Zeira (1993) "Income distribution and macroeconomics," *Review of Economic Studies* 60(1), 35–52, January

Hansen, Lars Peter (1982) "Large sample properties of generalized method of moments estimators," *Econometrica* 50(4), 1029–1054, July

Heston, Alan, and Robert Summers (1999) "The world distribution of income: A synthesis of intercountry and intracountry data to measure worldwide material well-being," Working Paper, Economics Department, University of Pennsylvania, August

Jones, Charles I. (1997) "On the evolution of the world income distribution," *Journal of Economic Perspectives* 11(3), 19–36, Summer

Kuznets, Simon (1955) "Economic growth and income inequality," *American Economic Review* 45(1), 1–28, March

Li, Hongyi, Lyn Squire, and Hengfu Zou (1998) "Explaining international and intertemporal variations in income inequality," *Economic Journal* 108, 1–18, January

Manski, Charles F. (1988) *Analog Estimation Methods in Econometrics* (London: Chapman and Hall)

Milanovic, Branko (1999) "True world income distribution, 1988 and 1993: First calculation based on household surveys alone," Working Paper, The World Bank, Washington DC, May

Perotti, Roberto (1996) "Growth, income distribution, and democracy: What the data say," *Journal of Economic Growth* 1(2), 149–187, June

Persson, Torsten, and Guido Tabellini (1994) "Is inequality harmful for growth?," *American Economic Review* 84(3), 600–621, June

Pritchett, Lant (1997) "Divergence, big time," *Journal of Economic Perspectives* 11(3), 3–17, Summer

Quah, Danny (1993) "Empirical cross-section dynamics in economic growth," *European Economic Review* 37(2/3), 426–434, April

— (1997) "Empirics for growth and distribution: Polarization, stratification, and convergence clubs," *Journal of Economic Growth* 2(1), 27–59, March

Ravallion, Martin (1997) "Can high-inequality developing countries escape absolute poverty?," *Economics Letters* 56(1), 51–57

Ravallion, Martin, and Shaohua Chen (1997) "What can new survey data tell us about recent changes in distribution and poverty?," *World Bank Economic Review* 11(2), 357–382, June

Ray, Debraj (1998) *Development Economics* (Princeton University Press)

Schultz, T. Paul (1998) "Inequality in the distribution of personal income in the world: How it is changing and why," *Journal of Population Economics* 11, 307–344

Silverman, Bernard W. (1981) "Using kernel density estimates to investigate multimodality," *Journal of the Royal Statistical Society, Series B* 43(1), 97–99

Snower, Dennis J. (1998) "Causes of changing earnings inequality," In "Income Inequality: Issues and Policy Options" Federal Reserve Bank of Kansas City pp. 69–133

Wolfson, Michael (1994) "Diverging inequalities," *American Economic Review* 84(2), 353–358, May

World Bank (1990) *World Development Report: Poverty* (Oxford University Press)