Discussion Paper No. 2002/38

One Third of the World's Growth and Inequality

Danny Quah*

March 2002

Abstract

This paper studies growth and inequality in China and India—two economies that account for a third of the world's population. By modelling growth and inequality as components in a joint stochastic process, the paper calibrates the impact each has on different welfare indicators and on the personal income distribution across the joint population of the two countries. For personal income inequalities in a China-India universe, the forces assuming first-order importance are macroeconomic: growing average incomes dominate all else. The relation between aggregate economic growth and within-country inequality is insignificant for inequality dynamics.

Keywords: China, distribution dynamics, Gini coefficient, headcount index, India, poverty, world individual income distribution

JEL classification: D30, O10, O57

---

*London School of Economics

This study is linked to the theme of Growth and Poverty discussed at the WIDER Development Conference, 25-26 May 2001, Helsinki. UNU/WIDER gratefully acknowledges the financial support from the governments of Denmark, Finland and Norway to the 2000-1 research programme

Acknowledgements

I thank the MacArthur Foundation for financial support. I have received many helpful suggestions from colleagues, including Abhijit Banerjee, Tim Besley, Richard Blundell, Andrew Chesher, Frank Cowell, Bill Easterly, Theo Eicher, Raquel Fernandez, Chico Ferreira, James Feyrer, Oded Galor, Cecilia García-Peñalosa, Louise Keely, and Branko Milanovic. Also useful were Clarissa Yeap's research assistance in the early stages of this work and Claudia Biancotti's Ec473 LSE seminar presentation in the spring of 2001. All calculations and graphs were produced using LATEX and the author's econometrics shell TSRF.

UNU World Institute for Development Economics Research (UNU/WIDER)
Katajanokanlaituri 6 B, 00160 Helsinki, Finland

Camera-ready typescript prepared by Danny Quah and Lorraine Telfer-Taivainen at UNU/WIDER
Printed at UNU/WIDER, Helsinki

ISSN 1609-5774
ISBN 92-9190-198-9 (printed publication)
ISBN 92-9190-199-7 (internet publication)

One third of the world's growth and inequality
by
Danny Quah *
LSE Economics Department
March 2002

1 Introduction

Three concerns underly all research on income inequality and economic growth. First, inequality might be causal for growth, raising or lowering an economy's growth rate. Understanding the mechanism then becomes paramount. How do alternative structures of political economy and taxation matter for this relation between inequality and growth? Does income inequality increase the rate of capital investment and therefore growth? Do credit and capital market imperfections magnify potentially adverse impacts of inequality, thereby worsening economic performance and growth? For concreteness, I will refer to this circle of related questions as the *mechanism* concern.

Second, even as economic growth occurs, the simultaneous rise in inequality—sometimes hypothesized, other times asserted—might be so steep that the very poor suffer a decline in their incomes. This is one of a set of beliefs underlying the anti-capitalism, anti-globalization, anti-growth movement. Although, not exhaustively descriptive, *anti-globalization* is the term I will use to refer to this second concern.

* I thank the MacArthur Foundation for financial support. I have received many helpful suggestions from colleagues, including Abhijit Banerjee, Tim Besley, Richard Blundell, Andrew Chesher, Frank Cowell, Bill Easterly, Theo Eicher, Raquel Fernandez, Chico Ferreira, James Feyrer, Oded Galor, Cecilia García-Peñalosa, Louise Keely, and Branko Milanovic. Also useful were Clarissa Yeap's research assistance in the early stages of this work and Claudia Biancotti's Ec473 LSE seminar presentation, in the spring of 2001. All calculations and graphs were produced using LaTeX and the author's econometrics shell tSrF.

Third is an all-else category of analyses that fall outside the first two. This incorporates concerns such as envy, equity, risk, peer group effects, or the economics of superstars (where the distribution of outcomes turns out more skewed than that of the important underlying characteristics). Thus, this category includes the more traditional motivations in research on income distribution and inequality, but that have become less emphasized in recent research that focus more on the mechanism and anti-globalization concerns.

This paper is part of a body of research that argues that the mechanism and anti-globalization concerns are empirically untenable. It seeks to sharpen the general points made in Quah (2001b) by concentrating on the world's two most populous nation states, China and India—only two points in a cross-country analysis, but fully one-third of the world's population.

The remainder of this paper is organized as follows. Section 2 describes related literature, and Section 3 develops the class of probability models underlying the approach in this paper. Sections 4–5 present the empirical results. Section 6 concludes. The technical appendix, Section 7, contains details on the estimation and data.

2 Related literature

A conventional wisdom recently emerged from empirical research on inequality and growth is how fragile empirical findings are, varying with auxiliary conditioning information, functional form specification, assumed patterns of causality, and so on (e.g., Banerjee and Duflo, 2000).

This state of affairs is unlike that at the beginnings of the subject. Then, Kuznets (1955) had asked if personal income inequality increased or declined in the course of economic growth. He documented both: looking across countries, from poorest to richest, within-country income inequality first rose and then fell.

Since most of the work there entailed defining and collecting data, it was painstaking and laborious. By contrast, modern researchers now using readily-available observations on growth and inequality can easily and routinely re-examine Kuznets's inverted U-shaped curve (e.g., Deininger and Squire, 1998). Interest therefore has shifted to more subtle issues: causality and mechanisms relating inequality and growth—see, e.g., Aghion, Caroli and García-Peñalosa (1999), Bénabou (1996), Galor and Zeira (1993), and the literature surveyed in Bertola (1999).

On these more complex questions, however, the data have given a less clearcut message. Results have varied, depending on auxiliary conditioning information and econometric technique. For instance, Alesina and Rodrik (1994), Perotti (1996), and Persson and Tabellini (1994) concluded that inequality and growth are negatively related, while Barro (2000), Forbes (2000), and Li and Zou (1998) reported a positive or varying relation. To some researchers, the situation has seemed so bad that they have simply concluded the data are not informative for interesting issues in inequality and growth, and have attempted to explain why this is so, within a particular model of inequality and growth (e.g., Banerjee and Duflo, 2000). In this view the data are noisy.

This paper takes no explicit stance on causality between inequality and growth, nor on the functional form relating them. Instead, it models inequality and growth jointly as part of a vector stochastic process, and calibrates the impact each has on a range of welfare indicators and on the individual income distributions—first within China and India and then taking the two countries together. The paper addresses simpler questions than those treated in the ambitious work attempting to trace out causality across growth and inequality.

This paper asks, when growth occurs, how do the poor fare? What difference have the historical dynamics of inequality and growth made for the incomes of one third of the world's population? If inequality were, indeed, to fall when growth is lower, does it fall enough to overcome the negative impact on the poor of slower economic growth overall? Alternatively, if within-country inequality were to rise, does that occur simultaneously with Chinese and India per capita incomes converging, so that overall individual income inequality across these two economies is falling?

Given the data extant, arithmetic alone suffices to retrieve useful answers to these questions. Here, the data are loud, not noisy: For a universe comprising China and India—one third of the world—for understanding the secular dynamics of personal incomes against a setting of cross-country inequalities, those forces of first-order importance are macroeconomic ones determining national patterns of growth and convergence. Rising average incomes dominate everything else. Within-country inequality dynamics are insignificant for determining inequality across people internationally.

Several earlier papers motivate my approach here. Deininger and Squire (1998) addressed questions closely related to those I pose above. They used regression analysis and more elaborate data, in contrast to the minimalist, arithmetic approach of this paper. They concluded, though, much the same as I do below: The poor benefit more from increasing aggregate growth by a range of factors, than from reducing inequality through redistribution. Deininger and Squire's view of growth and inequality as the joint outcome of some underlying, unobserved development process matches that in Section 3 below.

Dollar and Kraay (2001) studied directly average incomes of the poorest fifth of the population across many different economies. They noted those incomes rise proportionally with overall average incomes, for a wide range of factors generating economic growth. Put differently, it is difficult to find anything raising average incomes that doesn't also increase incomes for the very poor. They concluded, as I will below, that the poor benefit from aggregate economic growth, whatever is driving the latter. Similarly, Ravallion and Chen (1997) found in survey data that changes in inequality are orthogonal to changes in average living standards.

All these papers, in my interpretation, point to a consistent, quantitatively important characterization of the relation between growth and inequality. The characterization is one naturally viewed in terms of Fig. 2 below, and rounded out by the arithmetic calculations in this paper and in Quah (2001b).

More recent papers are related as well. Bourguignon (2001) performed calculations like those in Section 7.4.1 below. Sala-i-Martin (2002) uses within-country income shares as vector $\mathcal{I}$ (section 3 below), and therefore is able to specialize the calculations from what I give in sections 3 and 7. Although his techniques and emphases differ from those in this paper and in Quah (2001b), our motivations and conclusions are close, and complement each other's. Finally, Heston and Summers (1999), Milanovic (2002), and Sala-i-Martin (2002) constructed world income distributions—i.e., across over a hundred countries, not just those for the two I use here—but that similarly put together individual country statistics. Their methods, approaches, and data sources differ from mine, but the underlying ideas are the same.

Critics of this work have pointed out that income inequality statistics such as Gini coefficients, 90/10 ratios, mean/median income ratio, log standard deviations, and so on were never intended for examining the kinds of issues that I treat below nor for merging with the per capita income and population measures that I analyze. Thus, for instance, more detailed investigations into the very poor in any single country, made possible from surveys, field research or other individual-level microeconomic data, could well display tendencies different from those I derive—see, e.g., Atkinson and Brandolini (2001) or Dreze and Sen (1995). My calculations then, it is asserted, do no more than reveal the misleading nature of many income inequality statistics.

Perhaps so. However, it is also exactly these same statistics with which Bénabou (1996) begins his powerful and influential statement on how inequality matters importantly in economic growth—merging inequality and aggregate statistics, comparing Korea and the Philippines, in a similar spirit to what I do below. If measures like those I use mislead, then all such research is flawed (which I don't believe)—not just those studies, like the current one, that argue inequality is unimportant. The single set of data that we all use has different dimensions to it, and we cannot selectively ignore some and heed only others—thereby imposing biases based on whether certain conclusions seem a priori sensible.

3 Probability models for income distribution dynamics

Fix a country at a point in time, and let Y denote income, $\mathcal{I}$ a vector of income inequality measures, and F the distribution of Y across individuals. One entry in $\mathcal{I}$ might be the Gini coefficient; another might be the mean-median income ratio; yet a third might be the standard deviation of log incomes; further entries might be (within-country) income shares; and so on. Each element of $\mathcal{I}$ is a functional or a statistic of the distribution F. To emphasize that per capita income is the *arithmetic mean* or *expectation* of F, write it as $\mathcal{E}$. Economic growth is $\dot{\mathcal{E}}/\mathcal{E}$.

Asking about causality between growth and inequality is asking about the functions

$$\dot{\mathcal{E}}/\mathcal{E} = \phi(\mathcal{I}) \quad \text{or} \quad (\mathcal{I}) = \psi\left(\dot{\mathcal{E}}/\mathcal{E}\right)$$

(as in, e.g., Fig. 1). If that is the interest, econometric analysis can trace out ϕ and ψ.

By contrast, this paper models $\dot{\mathcal{E}}/\mathcal{E}$ and $\mathcal{I}$ jointly, taking them to be elements of equal standing in a vector stochastic process Z. Let Z_0 denote the vector of other variables in the system, including population P, so that

$$\{Z(t) : t \geq 0\}, \text{ with } Z \stackrel{\text{def}}{=} \begin{pmatrix} \dot{\mathcal{E}}/\mathcal{E} \\ \mathcal{I} \\ Z_0 \end{pmatrix}$$

constitutes the object to investigate. The current study can be viewed as describing an unrestricted vector autoregression in Z; it makes no assumptions on causality relations across the different entries of Z. The law of motion describing the dynamics of the income distribution F implies a law of motion for the vector Z. Conversely, when Z_0 is sufficiently extensive, Z's dynamics imply F's; when Z_0 is not complete, Z's dynamics restrict but do not fully specify F's.

Fig. 2 illustrates this. The right side shows the density f corresponding to the distribution F, at two time points t_0 and t_1, with the dashed line indicating f at t_0, the earlier time, and the solid line

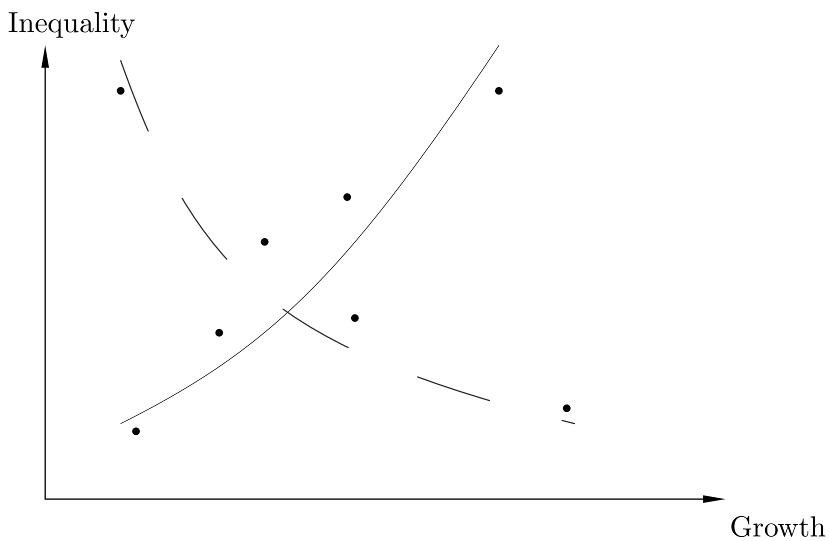

Fig. 1: Inequality and growth Does one systematically co-move with the other? Does one cause the other?

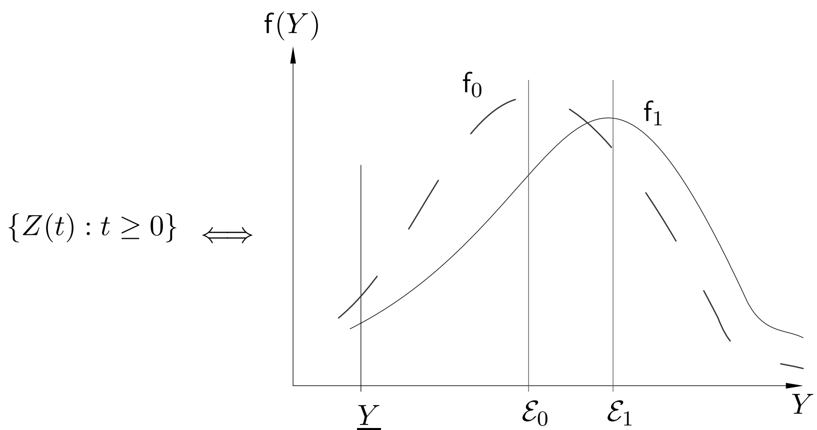

Fig. 2: Income distribution dynamics Vector Z's law of motion implies and is implied by income distribution dynamics. Densities f_0 and f_1 are for times t_0 and t_1 respectively, with $t_0 < t_1$.

indicating that at timepoint t_1. Associated with f_0 is its mean $\mathcal{E}_0$; similarly, associated with f_1 is its mean $\mathcal{E}_1$. Fig. 2 has, as an illustration, $\mathcal{E}_1 > \mathcal{E}_0$ so that economic growth, as measured by national income statistics, has occurred. If $\underline{Y}$ is some arbitrary but fixed income level, we can estimate the fraction of the population that remains with income below $\underline{Y}$ by calculating $\int_{Y \leq \underline{Y}} \mathsf{f}(Y)\, dY$ from knowledge of f, from time t_0 to time t_1. If we know the population P as well, then we can use this calculation to estimate the number of people living at incomes less than $\underline{Y}$. In the inequality literature, the statistic $\int_{Y \leq \underline{Y}} \mathsf{f}(Y)\, dY$ is sometimes called the *poverty headcount index*, and written $HC_{\underline{Y}}$; while the size of the population with incomes at most $\underline{Y}$ is written $P_{\underline{Y}}$ (e.g., equations (14) and (15) in Section 7 below).

The problem is we typically have only incomplete information on Z and f. But we can use knowledge on Z to infer restrictions on f, and then estimate statistics of interest like $HC_{\underline{Y}}$ and $P_{\underline{Y}}$.

To illustrate the idea, suppose F were assumed (or otherwise inferred) to be Pareto, and known up to the two parameters θ_1 and θ_2:

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

What restrictions does knowledge of Z imply for F? Equation (1) gives per capita income $\mathcal{E}$ and Gini coefficient $\mathcal{I}_G$ as

$$\mathcal{E} \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} y\, d\mathsf{F}(y) = (\theta_2 - 1)^{-1}\theta_2\theta_1,$$
$$\mathcal{I}_G \stackrel{\text{def}}{=} \left[2^{-1}\mathcal{E}(\mathsf{F})\right]^{-1} \int_{-\infty}^{\infty} \left(\mathsf{F}(y) - \frac{1}{2}\right) y\, d\mathsf{F}(y) = (2\theta_2 - 1)^{-1},$$

so that knowledge of the first two entries of Z alone gives

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

With more information in Z, the researcher can either estimate θ more precisely, using a method of moments technique as described in Section 7.1 below, or alternatively, relax the Pareto assumption for F. In either case, $HC_{\underline{Y}}$ and $P_{\underline{Y}}$ can then be straighforwardly estimated.

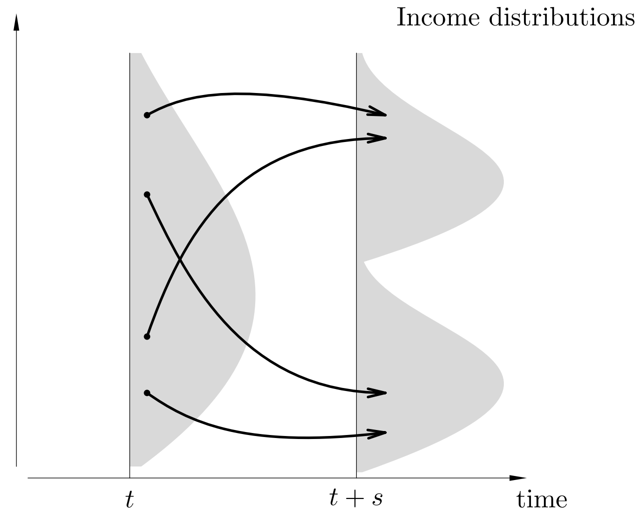

Fig. 3: Emerging twin peaks Cross-country per capita income distribution. Arrows show countries transiting across different parts of the cross section distribution.

Putting together the implied F's for different countries in the world would allow mapping the worldwide income distribution (Quah, 2001b). To appreciate the value of doing this, consider Fig. 3, which shows what has sometimes been referred to as an *emerging twin peaks* in cross-country income distribution dynamics (e.g., Quah, 2001a). This twin-peaks characterization, as many others in the macroeconomic growth literature, takes a country as a unit of observation. Thus, countries as large as China are treated the same way as those as small as Singapore, and income distributions *within* countries are ignored—the analysis takes everyone in the economy to have the same (per capita, average) income.

Information on income distributions within a country allow enriching the picture in Fig. 3 to something like Fig. 4. The black dots at time t indicate the per capita incomes of two hypothetical countries, with the darker shaded area around each depicting within-country individual income distributions. Thus, even as national per

Fig. 4: Individual income distributions Distribution dynamics within the emerging twin-peaks law that describes cross-country per capita income cross sections.

capita incomes evolve according to an emergent twin-peaks dynamic law, the distribution of incomes across people, within and across countries, can evolve and overlap in intricate ways.

4 China and India

China and India—although only two countries out of over a hundred in Fig. 3—carry within them a third of the world's population. They thus provide substantial insight into the dynamics in Fig. 4.

Table 1 records that between 1980 and 1992 China's per capita income grew from US$972 to US$1493, an annual growth rate of 3.58%. Over this period, India grew at a lower annual rate of 3.12%, increasing its per capita income from US$882 to US$1282. (Per capita incomes are purchasing power parity adjusted real GDP per capita in constant dollars at 1985 prices, series `rgdpch` from Summers and

	Per capita incomes (US$)			Population ($\times 10^6$)	
	1980	1992	$\dot{\mathcal{E}}/\mathcal{E}$	1980	1992
China	972	1493	3.58%	981	1162
India	882	1282	3.12%	687	884
US	15295	17945	1.33%	228	255

Table 1: Aggregate income and population dynamics. China, India, and the US

	Gini coefficient $\mathcal{I}_G$		
	1980	1992	Min. (year)
China	0.32	0.38	0.26 (1984)
India	0.32	0.32	0.30 (1990)
US	0.35	0.38	0.35 (1982)

Table 2: Inequality in China, India, and the US, by Gini coefficient (from Deininger and Squire, 1996)

Heston (1991).) For comparison, Table 1 also contains a row for the US, showing its 1.33% annual growth over 1980–1992, taking per US capita income from US$15295 to US$17945. The last two columns of Table 1 contain population figures, again from Summers and Heston (1991). By 1992, China had grown to over 1.1 billion people, with India approaching 0.9 billion.

It has been remarked many times elsewhere that China's fast-increasing per capita income came together with rises in inequality. Table 2 shows Gini coefficients for the same three countries, China, India, and the US, over 1980–1992. Inequality in China, as measured by the Gini coefficient, increased from 0.32 to 0.38, while that in India remained constant at 0.32. While the last column of the Table shows the increase in China's inequality is not monotone and India's inequality was not constant throughout—China had its low Gini coefficient of 0.26 over this period in 1984, while India's low was 0.30 in 1990—it is tempting to conclude from looking cross-sectionally at China and India that a fast-growing economy also has its inequality rise rapidly, while a slower-growing economy can keep inequality in

	$\underline{Y} = 2$; $HC_{\underline{Y}}$ $(P_{\underline{Y}}, 10^6)$	
	1980	1992
China	0.37–0.54 (360–530)	0.14–0.17 (158–192)
India	0.48–0.62 (326–426)	0.12–0.19 (110–166)

Table 3: Fraction of population and number of people with incomes less than US$2 per day. The range of estimates spans the different distributional assumptions described in Section 7.4.

check.

But what do Tables 1 and 2 imply for, say, the number of people in China and India living below a specific fixed income level? How rapidly were people exiting low-income states, given aggregate growth and actual changes in measured inequality? If, counterfactually, inequality had remained unchanged, how would aggregate growth alone have changed conditions for the poor? Or, again counterfactually, how much would inequality have had to increase, for the poor not to have benefited at all from aggregate growth?

Tables 3 and 4 provide answers to these questions, obtained using the calculations detailed below in Section 7. First, from the actual historical record in per capita income growth ($\dot{\mathcal{E}}/\mathcal{E}$), population, and Gini coefficients, we can, with weak additional assumptions on the parametric form of density f, work out how the entire distribution of individual income shifted between 1980 and 1992. Table 3 shows how the situation for the very poor changed over this time period. The fraction of the population living on less than US$2 per day (US$730 annually) varied from 0.37 to 0.54 in China in 1980; this corresponded to between 360m to 530m people.[1] By 1992, the fraction of population in that income range had fallen to 0.14 to 0.17, implying only between 158m to 192m people, given the population size then. In other words, over 1980–1992 China reduced the population in this very poor income range by between 210m to 338m people, even as

[1] Each entry in the Table is a range rather than just a single number since alternative distributional assumptions can be used in the calculation; see Section 7.4 below.

	$\mathcal{I}_G$, P constant: $-\dot{P}_{\underline{Y}}$	$HC_{\underline{Y}}$ constant: $\dot{\mathcal{I}}_G/\mathcal{I}_G$
China	33m/year	8.3%/year
India	17m/year	8.8%/year

Table 4: From 1980 perspective: Given aggregate growth, reduction in numbers of poor if inequality unchanged, and proportional inequality increase per year to maintain poverty numbers

inequality and total population rose.

The situation for India is less surprising, as measured inequality there remained constant, and only aggregate economic growth occurred.[2] But since the total Indian population also rose, it might well have been that the poor did increase in number. Table 3 shows that did not happen. Between 1980 and 1992, the number of Indians living on less than US$2 a day fell from 326m–426m to less than half that, 110m–166m. The fraction of the Indian population in this income range fell from approximately half to perhaps one-fifth, likely less. India reduced the population living in the very poor income range by about a quarter of a billion, a number comparable to the change in China.

If the world comprised only China and India put together, it would show a number of interesting features. First, the country that grew faster on aggregate also had inequality rise more—the upward-sloping schedule in Fig. 1 is that that is relevant. Second, however, even despite this positive relation between growth and inequality, overall the world's poor benefited dramatically from economic growth. Over the course of little more than a decade, about half a billion people—out of a total population across the two countries of about 1.6 to 1.9 billion—exited the state of extreme poverty. This decline in sheer numbers of the very poor divided about equally between China and India.

Table 4 takes the argument further. Suppose population were held

[2] See again, however, the discussion at the end of section 2 and the more detailed picture available from studies such as Dreze and Sen (1995).

constant at 1980 levels in China and India. The left panel in the table shows that if inequality too were held constant at its 1980 levels, then aggregate growth alone would have removed from being very poor 33m people a year in China, and 17m people a year in India. The right panel shows that to keep constant the number of people living on incomes less than US$2 a day as aggregate growth proceeded, the Gini coefficient would have had to rise at a proportional growth rate of 8.3% per year in China, and 8.8% per year in India. Such rapid and large increases in inequality are unprecedented in world history.[3] China's increase in inequality would have had to more than double, and be kept at that rate for a dozen years, to nullify the beneficial effects of its high aggregate growth rate.

I conclude from the discussion here that, given the historical experience in China and India, aggregate economic growth might well come about only with increases in inequality. However, given magnitudes that are historically reasonable, growth is unambiguously beneficial—especially for the poor in general, and even for the poor in particular when inequality rises.

5 Extensions

This section expands on the discussion in Section 4 above. It provides further quantification and illustration to the conclusions there.

[3] The only possible exceptions are the transition economies and Russia after the collapse of the Soviet Union: see, e.g., Ivaschenko (2001) and Shorrocks and Kolenikov (2001). The seven instances that Li, Squire and Zou (1998) identified with statistically and quantitatively significant time trends in Gini coefficients only saw proportional growth rates of 1.02% (Australia), 1.04% (Chile), 3.18% (China), −1.71% (France), −1.18% (Italy), 1.61% (New Zealand), and 1.46% (Poland) [this author's calculations, from Table 4 in Li, Squire and Zou (1998)] taken linearly over a single time period—multiple time periods would imply yet smaller growth rates.

Deininger and Squire (1996), Li, Squire and Zou (1998), and Quah (2001b) have presented calculations formalizing how most of the variation in measured inequality is *across* countries. Inequality changes hardly at all in time. It is not that inequality—for physical reasons or otherwise—cannot vary much; it is that the workings of economies lead to inequality hardly changing through time. Honduras's 1968 income inequality of 62% (Gini coefficient) is 2.4 times that of Belgium's 26% in 1985. But at the same time, over the entire post-War era, income inequality in Belgium never rose above 28% while Honduras's never fell below 50%.

This fact has implications for panel data analyses of inequality and growth (Quah, 2001b). Panel data econometric methods that condition out individual effects—almost all do—end up removing all the important variation in inequality data. Moreover, since economic growth has its principal variation in the orthogonal direction—in time rather than across countries (e.g., Easterly, Kremer, Pritchett and Summers, 1993)—panel data methods shoehorn inequality and growth data into spuriously better fit with each other. In other words, econometric methods that condition out individual effects represent here a methodologically-suspect attempt to remove statistical biases. That justification has often been simply imported from other fields of economics, without due attention to why applying such techniques to study inequality and growth might be inappropriate.

From Table 3 we concluded that China and India together reduced the number of people living on less than US$2 a day by between 36m and 50m each year. Fig. 5 graphs this same reduction from 1980 to 1992 in the number of people with incomes less than US$2 per day, given the actual historical outcomes in income inequality and economic growth. China and India alone shifted 508m people, more than 12% of the world's population, about one-third of the world's then-poor, out of poverty.

Fig. 6 shows, for each economy, the amount of poverty reduction per year that would have occurred from aggregate growth alone, had inequality and population size remained constant at their 1980 values. Obviously, the faster is economic growth, the faster would $P_{\underline{Y}}$ fall. The Fig. emphasizes that to reduce poverty worldwide it is in the very

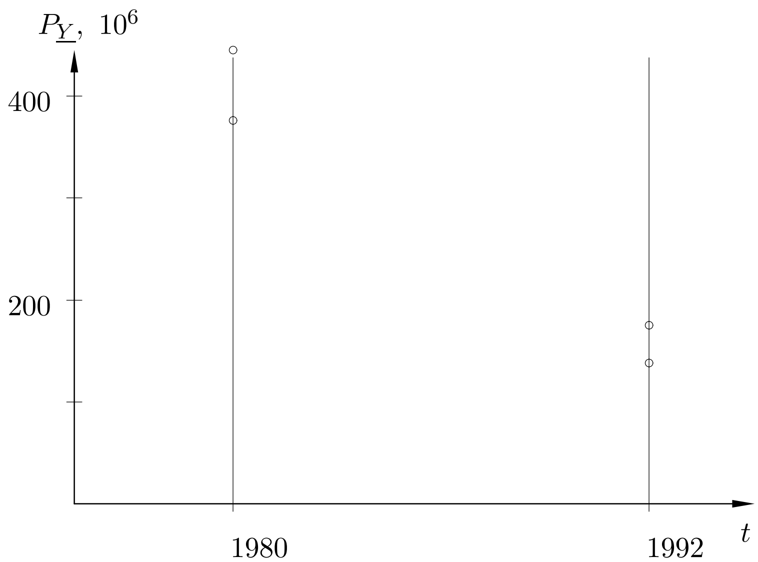

Fig. 5: Estimated absolute poverty reduction Each country is a single dot, one each for 1980 and for 1992—with height equal to the average of $P_{\underline{Y}}$ across the distributional assumptions for F described in Section 7.4. The sum total across China and India in 1980 is 821m; in 1992, 313m.

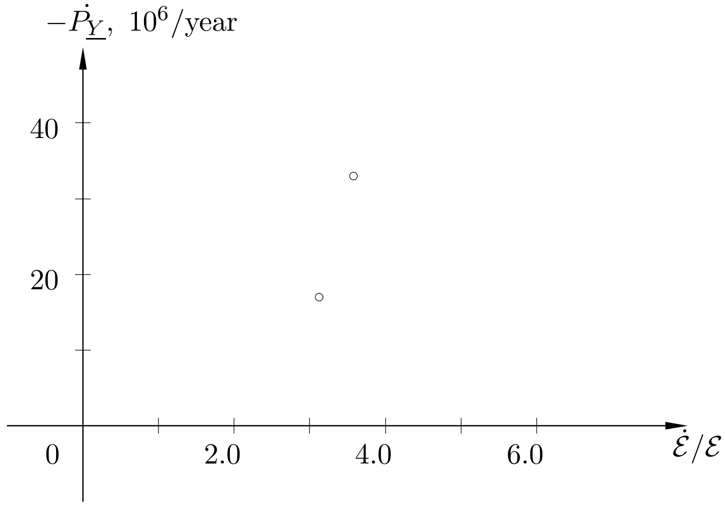

Fig. 6: Growth alone Each country is a single dot, with height equal to the arithmetic average of $-\dot{P}_{\underline{Y}}$, 10^6/year across distributional assumptions for F described in Section 7.4.

large economies like China and India where high growth is needed.

Fig. 7 shows, again for each economy, the proportional growth rate of inequality that would be required to nullify the benefits of growth, had population remained constant at its 1980 value. The dark dots towards the bottom of the Fig. shows the actual growth rate in inequality that occurred over the sample. (These numbers would be lower than those in footnote 3, for the multiple/single timeperiod reason described there.) Counterfactual increases in inequality of the magnitude that would be needed to overcome the poverty-reducing impact of economic growth—the upper part of the picture, in light dots—are far outside the range of historical realization.

Quah (2001b) presents counterparts of Figs. 5–7 for the 100 or so countries for which data are available. The conclusions above are reinforced; unsurprisingly, China and India dominate the picture.

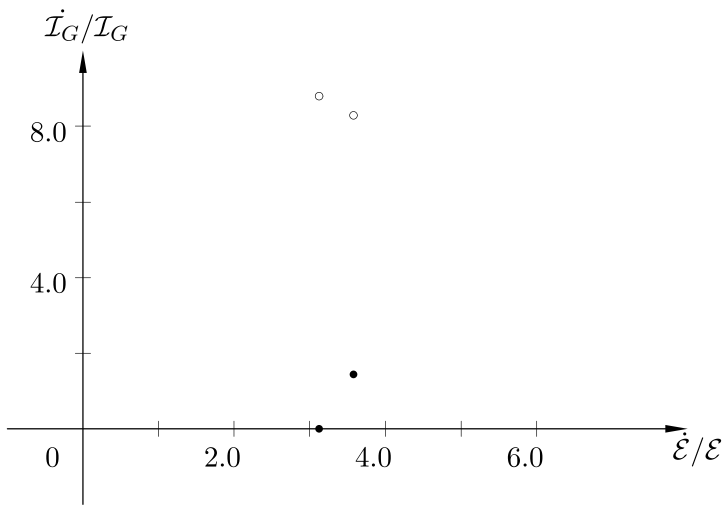

Fig. 7: Inequality to nullify growth Each country is a single dot. The light dots have height equal to the arithmetic average of $\dot{\mathcal{I}}_G/\mathcal{I}_G$ across distributional assumptions for F described in Section 7.4. The dark dots show actual $\dot{\mathcal{I}}_G/\mathcal{I}_G$ realized.

6 Conclusions

Much recent research on inequality and growth has taken one of two possible approaches: The first is explicit, and that is to see if inequality causes growth. The second, typically left implicit, is to see if, even as growth occurs, the poor might be disadvantaged anyway, because inequality has risen so dramatically. This paper has shown that for China and India—only two points in a cross section but one third of the world's population—neither of these possibilities is empirically tenable.

More traditional motivations—risk, poverty, equity—for studying inequality, however, remain. Indeed, they are reinforced as worldwide inequality continues to evolve, driven by powerful economic forces. But these motivations have little to do with the more recent analyses of the relation between inequality and growth.

This paper has applied a simple arithmetic approach to obtain its findings. It has asked, given historical patterns of growth and inequality, how have income distributions within the each of and across China and India evolved? How have the poor fared in the two countries as per capita incomes and rich-poor differences have changed? Whether the growth and inequality data are unable to speak clearly on questions of causality or whether they only imply weak empirical relations, the data are unequivocal on the questions I posed. The data are loud, not noisy.

The principal finding of the paper is two-fold: First, only under inconceivably high increases in inequality would economic growth not benefit the poor. The magnitudes of improvement in living standards due to aggregate economic growth simply overwhelm any putative deterioration due to increases in inequality. Second, any mechanism where inequality causes economic growth, positively or negatively, is empirically irrelevant for determining outcomes for individual income distributions. I have obtained these results for China and India in this paper. A complete study—rounding out the remaining two-thirds left over from the title of this paper—is in Quah (2001b).

This paper has taken care to assume no single view on the causal mechanisms relating inequality and growth. It has pointed out that whatever it is that drives economic growth in the large, those forces—be they macroeconomic, technological, political, or institutional—are dramatically important for improving the lot of the poor when they lead to economic growth. And similarly so in the opposite direction when they lead to economic stagnation.

What might seem an appealing possibility to raise here is that income inequality could, positively or negatively, truly cause aggregate economic growth, so that this paper's principal finding would then only reinforce the importance of distributional and inequality concerns over macroeconomic growth. However, even in reduced-form regressions of growth on inequality, the R^2 fit can never be very high: The directions of principal variation in the two variables are just too different (Quah, 2001b). Therefore, even in the best of circumstances, even with no ambiguity on the direction of causality, many other factors beyond inequality influence economic growth. And *all* of them, through their impact on the aggregate income level, affect the poor—independently of inequality's effect on economic growth.

Finally, it is not a telling criticism of the work in this paper to say that because it uses Gini coefficients or other standard inequality measures, it does not get at the true nature of inequality, whatever that might mean. Because the paper's methods characterize the entire income distribution, the focus on specific inequality measures is only for practicaly convenience, not conceptual necessity. However, precisely the same measures are used in all other studies of growth and inequality I know—in particular, in the many regression studies that claim to show one causal relation or another between these two quantities. If the data used are inappropriate here, then they are similarly so there too. Indeed, one might view the calculations here as simply taking a logical step prior to other work in its drawing out an interpretation to the indexes used in studies of inequality and growth.

References

Aghion, Philippe, Eve Caroli, and Cecilia García-Peñalosa (1999) "Inequality and economic growth: The perspective of the new growth theories," *Journal of Economic Literature* 37(4), 1615–1660, December

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

Atkinson, Anthony B., and Andrea Brandolini (2001) "Promise and pitfalls in the use of secondary datasets: Income inequality in OECD countries as a case study," *Journal of Economic Literature* 39(3), 771–799, September

Banerjee, Abhijit V., and Esther Duflo (2000) "Inequality and growth: What can the data say?," Working Paper, Economics Department, MIT, May

Barro, Robert J. (2000) "Inequality and growth in a panel of countries," *Journal of Economic Growth* 5(1), 5–32, March

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

Bertola, Giuseppe (1999) "Macroeconomics of distribution and growth," In *Handbook of Income Distribution,* ed. Anthony B. Atkinson and Francois Bourguignon, vol. 1 (North Holland Elsevier Science) chapter 9, pp. 477–540

Bourguignon, François (2001) "The growth elasticity of poverty reduction: Explaining heterogeneity across countries and time periods," Working Paper, Delta and the World Bank, Paris, December

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

——— (1998) "New ways of looking at old issues: Inequality and growth," *Journal of Development Economics* 57, 259–287

——— (2002) "Revisiting inequality: New data, new results," Working Paper, The World Bank, Washington DC, March

Dollar, David, and Art Kraay (2001) "Growth is good for the poor," Working Paper, DRG, The World Bank, Washington DC, March

Dreze, Jean, and Amartya Sen (1995) *India: Economic Development and Social Opportunity* (Oxford: Oxford University Press)

Easterly, William, Michael Kremer, Lant Pritchett, and Lawrence H. Summers (1993) "Good policy or good luck? Country growth performance and temporary shocks," *Journal of Monetary Economics* 32(3), 459–483

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

Forbes, Kristin J. (2000) "A reassessment of the relationship between inequality and growth," *American Economic Review* 90(5), 869–887, December

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, University of Pennsylvania, Philadelphia, August

Ivaschenko, Oleksiy (2001) "Growth and inequality: Evidence from transitional economies," Working Paper, Gothenburg University, December

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

Li, Hongyi, and Hengfu Zou (1998) "Income inequality is not harmful for growth: Theory and evidence," *Review of Development Economics* 2, 318–334

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 (2002) "True world income distribution, 1988 and 1993: First calculation based on household surveys alone," *Economic Journal* 112(476), 51–92, January

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

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

——— (2001a) "Cross-country growth comparison: Theory to empirics," In *Advances in Macroeconomic Theory,* ed. Jacques Dreze, vol. 133 of *Proceedings of the Twelfth World Congress of the International Economic Association, Buenos Aires* (London: Palgrave) chapter 16, pp. 332–351

——— (2001b) "Some simple arithmetic on how income inequality and economic growth matter," Working Paper, Economics Department, LSE, London, June. Given as the first *Journal of Applied Econometrics Lecture* at the Econometric Society Meetings, Auckland NZ, July 2001

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

Sala-i-Martin, Xavier (2002) "The disturbing "rise" of world inequality," Working Paper, Columbia University, New York, March

Shorrocks, Anthony F., and Stanislav Kolenikov (2001) "Poverty trends in Russia during the transition," Working Paper, UNU/WIDER, Helsinki, May

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

Summers, Robert, and Alan Heston (1991) "The Penn World Table (Mark 5): An expanded set of international comparisons, 1950–1988," *Quarterly Journal of Economics* 106(2), 327–368, May

UNU (2000) *UNU/WIDER-UNDP World Income Inequality Database,* v 1.0 ed. (United Nations University/WIDER)

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

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

7 Technical Appendix

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.

I develop here an alternative empirical framework that is general, flexible, and convenient. The approach 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.

I 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, 2002). 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

$$F_{W,t}(y) = P_{W,t}^{-1} \sum_{j=1}^{N} F_{j,t}(y) \times P_{j,t}, \quad y \in (0, \infty) \tag{2}$$

with the world population

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

Differentiating (2) with respect to y gives the implied density for the worldwide distribution of income as the weighted average of individual country income distribution densities:

$$f_{W,t}(y) = \sum_{j=1}^{N} f_{j,t}(y) \times (P_{j,t}/P_{W,t}), \quad y \in (0, \infty). \tag{3}$$

Knowing the distribution F_W means we can calculate directly all the inequality indexes we wish—whether or not particular indexes aggregate becomes irrelevant.

7.1 Estimating individual income distributions

Given the quantities on the right of equation (3) 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.

Write the vector of potentially-observable functionals as a collection

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

(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 $\varphi_t \in (\mathbf{B}, \mathfrak{B})$,

$$X_{l,t} = \mathbf{T}_l(\varphi_t, P_t), \quad l = 1, 2, \ldots , d. \tag{4}$$

Without loss or ambiguity, I will also write $\mathbf{T}_l(F_t, P_t)$ to denote the right hand side of (4). 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 $\theta_t \in \mathbb{R}^p$,

$$F_t = F(\cdot|\theta_t) \stackrel{\text{def}}{=} \mathsf{F}_{\theta_t}. \tag{5}$$

(In equation (5) 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 (5) 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

$$\hat{\theta}_t \stackrel{\text{def}}{=} \operatorname*{arg\,min}_{\theta \in \mathbb{R}^p} \;\; (\mathbf{T}(\mathsf{F}_\theta, P_t) - X_t)' \Omega \left(\mathbf{T}(\mathsf{F}_\theta, P_t) - X_t\right), \qquad \Omega \;\, d \times d \text{ positive definite.} \tag{6}$$

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

$$Q_{X_t}(\theta) = (\mathbf{T}(\mathsf{F}_\theta, P_t) - X_t)' \, \Omega \left(\mathbf{T}(\mathsf{F}_\theta, P_t) - X_t\right), \tag{7}$$

and denoting $\theta_{t,0}$ as the probability limit of (6), 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 (6) in (5) gives an estimator for F_t in each economy j. Plugging the result for each j in turn into (2)–(3) 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 7.4 below provides some explicit analytically worked-out examples of this procedure.

7.2 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 7.1 above—they are readily used in estimating and characterizing the distributions $F_{j,t}$. Conversely, if they are not observable but an estimate of $F_{j,t}$ is available, then estimates for $\mathbf{T}_l$ can, instead, be induced.

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

$$\mathcal{E}(\mathsf{F}, P) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} y \, d\mathsf{F}(y). \tag{8}$$

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

$$\mathcal{I}_G(\mathsf{F}, P) \stackrel{\text{def}}{=} \left[2^{-1}\mathcal{E}(\mathsf{F})\right]^{-1} \int_{-\infty}^{\infty} \left(\mathsf{F}(y) - \frac{1}{2}\right) y \, d\mathsf{F}(y) \tag{9}$$

(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,

$$Y_{0.2i}(\mathsf{F}, P) \stackrel{\text{def}}{=} \sup_{y \in \mathbb{R}} \{y \, | \mathsf{F}(y) \leq 0.2i\} \tag{10}$$

$$S_{0.2i}(\mathsf{F}) \stackrel{\text{def}}{=} \left(\int_{-\infty}^{Y_{0.2i}(\mathsf{F},P)} y \, d\mathsf{F}(y) \right) \times \mathcal{E}(\mathsf{F}, P)^{-1}. \tag{11}$$

The first of these, equation (10), defines the $(20 \times i)$-th percentile income level; the left-hand side is also known as the i-th quintile. The pair (10)–(11) generalizes to arbitrary percentile shares, but in practice the more general versions are rarely used (see, however, (12), (13), and (17) below).

Concepts (9)–(11) 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 (2002) has used household data to construct *within-decile average incomes* across many different countries. These fit within our framework as follows. Define

$$Y_{0.1i}(\mathsf{F}, P) \stackrel{\text{def}}{=} \sup_{y \in \mathbb{R}} \{y \,|\mathsf{F}(y) \leq 0.1i\}, \quad i = 0, 1, \ldots, 9, \tag{12}$$

and let

$$\begin{aligned} \mathcal{E}_{0.1i}(\mathsf{F}, P) &\stackrel{\text{def}}{=} \int_{Y_{0.1\times(i-1)}}^{Y_{0.1i}} y \, d\mathsf{F}(y), \quad i = 1, \ldots, 9, \\ \mathcal{E}_1(\mathsf{F}, P) &\stackrel{\text{def}}{=} \int_{Y_{0.9}}^{\infty} y \, d\mathsf{F}(y). \end{aligned} \tag{13}$$

Similar to (10) above, equation (12) 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 (2002) can thus be merged with that below if we use the decile averages $\mathcal{E}_{0.1i}$ from (13) (or even the deciles themselves $Y_{0.1i}$ in (12)) 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:

$$HC_{\underline{Y}}(\mathsf{F}, P) \stackrel{\text{def}}{=} \mathsf{F}(\underline{Y}) = \int_{-\infty}^{\underline{Y}} d\mathsf{F}(y). \tag{14}$$

Equation (14) gives a *poverty headcount index*, i.e., the fraction of population below a given income level $\underline{Y}$. Record also the absolute size of the population with those incomes:

$$P_{\underline{Y}}(\mathsf{F}, P) \stackrel{\text{def}}{=} P \times \mathsf{F}(\underline{Y}). \tag{15}$$

Finally, define:

$$PGI_{\underline{Y}}(\mathsf{F}, P) \stackrel{\text{def}}{=} \frac{\int_{-\infty}^{\underline{Y}} y \, d\mathsf{F}(y)}{\underline{Y}}. \tag{16}$$

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 (10) and let $Y_{0.5}$ denote the *median*

$$Y_{0.5}(\mathsf{F}, P) \stackrel{\text{def}}{=} \sup_{y \in \mathbb{R}} \left\{ y \,|\mathsf{F}(y) \leq \frac{1}{2} \right\}, \tag{17}$$

and then, using (8), (9), and (17), define a *polarization index*

$$Pz(\mathsf{F}, P) \stackrel{\text{def}}{=} \left[(1 - \mathcal{I}_G)\mathcal{E} - \frac{\int_{-\infty}^{Y_{0.5}} y \, d\mathsf{F}(y)}{\int_{-\infty}^{Y_{0.5}} d\mathsf{F}(y)} \right] \times \frac{2}{Y_{0.5}}. \tag{18}$$

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 (18).

All the functionals so far considered—apart from $P_{\underline{Y}}$ in (15)—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 (14), one might be interested in the rate of flow of people past the fixed income level $\underline{Y}$. This is

$$\begin{aligned} Fl_{\underline{Y}}(\mathsf{F}_{\theta_t}, P_t) &\stackrel{\text{def}}{=} -\frac{d}{dt} \left(\mathsf{F}_{\theta_t}(\underline{Y}) \cdot P_t \right) \qquad (19) \\ &= -\left[P_t \frac{d}{dt} \mathsf{F}_{\theta_t}(\underline{Y}) + \mathsf{F}_{\theta_t}(\underline{Y}) \frac{dP_t}{dt} \right]. \end{aligned}$$

Equation (19) 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 $\mathcal{I}_G$. I will use this simultaneous relationship below in sections 7.4.1 and 7.4.2. 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.

7.3 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 I organize all subsequent discussion. Equation (2) 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. I 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 (6); 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 (6)—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, 2002; Silverman, 1981). One should never need to construct any of (9)–(19), and go through (6), to characterize the distribution $F_{j,t}$. It is because such individual-level data are not readily available—instead statistical agencies have calculated and made available only different, aggregative statistics of the underlying data—that we are led to estimation by (6).

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 (6)–(7) 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.

7.4 Induced statistics and parametric examples

I 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:

$$\dot{\mathcal{E}}/\mathcal{E} = \xi > 0. \tag{20}$$

I 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 (19). We will be interested in the value of (19) 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 (19) 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

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

7.4.1 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 (8)–(11) of section 7.2 are:

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

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

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

If estimation (6) 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:

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

These can be used, in any case, as starting values in an iterative solution to (6). Heston and Summers (1999) used these as the estimates for their study.

Explicit formulas for some of the dynamics are then available:

$$\begin{aligned}
\dot{\mathcal{E}}/\mathcal{E} &= \dot{\theta}_1 + \frac{1}{2}\dot{\theta}_2,\\
\dot{\mathcal{I}_G}/\mathcal{I}_G &= \frac{\mathsf{f}_{\mathsf{N}(0,1)}([\theta_2/2]^{1/2})}{2\mathsf{F}_{\mathsf{N}(0,1}([\theta_2/2]^{1/2}) - 1}\cdot(\theta_2/2)^{1/2} \times \dot{\theta}_2/\theta_2,
\end{aligned}$$

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 4. Other distributional hypotheses will, as with the log Normal, require at least some of the results calculated numerically as closed-form expressions are intractable.)

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

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

so that for any fixed y,

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

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.

7.4.2 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 (8)–(11) of section 7.2 then are:

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

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 (6) is available when only $\mathcal{E}$ and $\mathcal{I}_G$ are observed:

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

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

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

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, require

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

or

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

To achieve this, we need

$$\dot{\mathcal{I}_G}/\mathcal{I}_G = \left(\frac{2\theta_2}{2\theta_2 - 1}\right)\left[\log\left(\theta_1/\underline{Y}\right) + (\theta_2 - 1)^{-1}\right]^{-1}\xi. \qquad (21)$$

Equation (21) 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 case 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.

7.5 Data

This paper merges data from Deininger and Squire (1996), Summers and Heston (1991), and UNU (2000). The updated and expanded inequality data in Deininger and Squire (2002) are not, as of this writing, yet distributed for general use.