Journal of Economic Literature 2008, 46:1, 57–94
http:www.aeaweb.org/articles.php?doi=10.1257/jel.46.1.57

What Do We Know about Global Income Inequality?

SUDHIR ANAND AND PAUL SEGAL*

In this paper, we review the recent literature on global interpersonal income inequality. While all estimates agree that the level is very high, with a Gini of between 0.630 and 0.686 in the 1990s, there is no consensus regarding the direction of change. We discuss methodological issues, including the use of national accounts versus survey-based estimates of mean income (or consumption) and the choice of purchasing power parity exchange rates. Findings of a rise or fall in global income inequality are not robust across different estimation methods and datasets. Given the diversity of estimates and various sources of uncertainty, including gaps and errors in the underlying data, we conclude there is insufficient evidence to determine the direction of change in global interpersonal inequality in recent decades.

1. *Introduction*

The last few years have seen a spate of papers estimating global income inequality. Their appearance is in part motivated by a desire to understand the effects of "globalization" and has been made possible by recent increases in the availability of data on income distributions within countries. Controversy centers on whether inequality has increased or decreased in the recent past. The direction and magnitude of change have been highly charged questions with some authors arguing that globalization has benefited the rich disproportionately, while others argue that it has reduced world income inequalities. Various findings are cited in the media, including the financial press, typically to support one or another position on globalization.

In this paper, we will review recent studies on global interpersonal inequality that, using different methods of estimation and different datasets, cover time periods up to the 1990s or later (1989 in the case of T. Paul Schultz 1998). They all estimate the level of and change in global interpersonal income or consumption inequality and use a variety of inequality measures. Earlier papers have also estimated global inequality, such as Albert Berry, François Bourguignon, and Christian Morrisson (1983) and Margaret E. Grosh and E. Wayne Nafziger (1986), but they were based on very limited income distribution data and, in this regard, the literature has advanced

* Anand: University of Oxford and Harvard University. Segal: University of Oxford. For helpful comments or discussion, we would like to thank Tony Atkinson, Lincoln Chen, Angus Deaton, James Foster, Camelia Minoiu, and Amartya Sen. We are also very grateful for the suggestions of the Editor and three anonymous referees. Anand gratefully acknowledges research support from the Rockefeller Foundation and the DKR/CAR Fund.

considerably.[1] Several other papers look solely at trends in within-country inequality around the world. For example, Giovanni Andrea Cornia and Sampsa Kiiski (2001) examine changes in income and consumption inequality within countries, and James K. Galbraith, Lu Jiaqing, and William A. Darity (1999) analyze the evolution of inequality of average earnings among branches of industry in sixty-six countries. Yet other studies, such as Glenn Firebaugh (1999, 2003) and Andrea Boltho and Gianni Toniolo (1999), estimate between-country inequality only, while Aren Melchior, Kjetil Telle, and Henrik Wiig (2000) estimate between-country inequality and report trends in regional Gini coefficients. However, none of these studies constructs a measure of global inequality that takes account of both between- and within-country inequality and are, therefore, outside the scope of this review.

A range of issues arise regarding the methods and data used by the studies under review and we examine in detail the relative merits of their estimation strategies. The studies differ in their manner of constructing a global distribution from the limited available data, particularly in their use of household surveys versus national accounts, their use of different purchasing power parity (PPP) exchange rates, and their estimation of within-country inequality. We discuss the implications of these differing methods and make recommendations regarding what we consider to be best practice. We also argue, more generally, that insufficient attention is typically given to both the choice of methodology and the uncertainty implied by the variable quality of data.

The changes in inequality found in these studies have often been adduced as evidence for or against the benefits of increased international economic integration. Quite apart from the problem of attributing causality,[2] we contend that the measured changes do not appear to be statistically significant on the basis of the standard errors estimated in some of the studies. Some changes, such as in Branko Milanovic (2002), appear large for the time period over which they are measured, but they are nonetheless small relative to plausible standard errors. Other sources of uncertainty (e.g., measurement and estimation problems) that are not incorporated in the estimated standard errors would lead to even wider confidence intervals. Such uncertainty, combined with the disagreement among the studies, leads us to the view that we cannot tell whether global inequality has increased or decreased in the recent past on the basis of existing findings.

The paper proceeds as follows. Section 2 begins by asking what kind of global inequality we want to measure and why. We summarize the key findings of the studies in section 3. In section 4, we discuss the methodological and data questions that arise and consider the possible biases implied by the different methods used. Section 5 critically reviews each study in the light of these methodological questions. Section 6 discusses the decomposition of global inequality into between-country and within-country components, and the significance of China and India. Section 7 turns to estimation errors and discusses how confident we should be of the various estimates and section 8 is in conclusion.

2. *What do We Want to Measure and Why?*

There are many reasons to be interested in global inequality. Three angles of interest,

[1] Berry, Bourguignon, and Morrisson (1983, p. 219) use data from "the developed countries and about forty less developed countries," stating that "for many L.D.C.s data at the national level is either non-existent or extremely weak." Even as late as 1992 income distribution estimates were available for only 41 out of 185 countries listed in the World Bank's (1992) *World Development Report*.

[2] For example, much of the increase in incomes of the poor in China occurred as a result of changes in government policy on domestic foodgrain prices in the early 1980s and mid 1990s, and had little if anything to do with increased international economic integration (Carl Riskin 2006).

ranging from the moral to the explanatory, can be readily identified. First, we may be interested in global inequality intrinsically, as large disparities in individuals' incomes may be considered unjust. Secondly, we may be interested in global inequality as an explanation for, or predictor of, some phenomenon of interest. Thus, unequal voting or bargaining power in international institutions may be a reflection of income inequalities among countries; or migration may be partly determined by global income inequalities as relatively poor people migrate to raise their living standards. Finally, we may be interested in global inequality as a predicted outcome of a theory, such as the convergence in per capita incomes across countries predicted by neoclassical growth theory or the divergence predicted by dependency theory.

The appropriate definition of "global inequality" depends on the purpose at hand. Milanovic (2005) provides a useful distinction between three concepts of world income inequality. Concept one is inequality among countries in their levels of average per capita income, with each country counting as a unit. Concept two is what we refer to as *between-country inequality*, which is inequality among individuals in the world with each individual assigned the average per capita income of his or her country of residence. Concept three, the focus of this review, is *global interpersonal inequality* or *global inequality* for short, which is inequality among individuals in the world with each individual assigned his or her own (per capita household) income. Concept two can readily be seen as the same as the "between-country component" of global inequality, measuring what global inequality would be if incomes were to be equally distributed among individuals within each country. Finally, to Milanovic's three concepts we would add a "concept zero inequality," which refers to inequality among countries ranked by total (not per capita) income. The population unit of concepts zero and one is the country, while that of concepts two and three is the individual.[3] In all cases, it remains to choose an appropriate "income" concept, e.g., consumption expenditure or income, assigned to the population unit in question (country or individual). This is important because, for instance, concept three inequality applied to income could move in a different direction from concept three inequality applied to consumption expenditure. This issue is discussed later.

To measure global inequality, we must also choose a set of exchange rates with which to convert national currencies into a common numeraire. The options are, broadly speaking, market exchange rates (say, relative to the U.S. dollar) versus PPP exchange rates.[4] PPP exchange rates take into account price differences across countries. They allow for the fact that a dollar's worth of rupees, bought on the currency markets, will buy more of most goods and services in India than the same dollar would buy in the United States. For developing countries, incomes measured using PPP exchange rates can be three or four times higher than when measured at market exchange rates.

Which definition of global inequality and which exchange rate are appropriate depend on the question being asked. Robert H. Wade (2001) suggests that market exchange rates are more appropriate than PPP exchange rates "for most of the issues that concern the world at large," including "migration flows" and "the extent of marginalization of developing countries in the world polity; and, more broadly, the economic and geopolitical impact of a country (or region) on the rest of the world." It seems more plausible to us that relative incomes measured at PPP exchange rates would be the better predictor

[3] Concepts two and three inequality could in principle also be defined across households rather than individuals.

[4] The two options are broad in the sense that choices remain within each: there will be different methods of smoothing market exchange rates over the year, and there are different methods for calculating PPP exchange rates. We consider PPP exchange rates in more detail later in the paper.

of migration flows. Remittances sent by migrants are indeed exchanged into national currencies on the market, but to the extent that people migrate to raise their own standard of living, it is PPP exchange rates that matter. On the other hand, market exchange rates do seem to be more appropriate in measuring "the economic and geopolitical impact of a country (or region) on the rest of the world." In this case, however, the variable of interest is presumably total income, not per capita income. This is one of the factors that underlies China's significance in world politics and makes India and Brazil important in international trade negotiations. The appropriate inequality concept here would seem to be concept zero inequality, i.e., gross national income across countries measured at market exchange rates.

Other questions call for different combinations of inequality concept and exchange rate. Consider the question of convergence between countries. This is based on concept one inequality where the country is the population unit, assigned its level of per capita income. As the variable of interest is the level of output per head, it is natural to use PPP exchange rates (e.g., Lant Pritchett 1997).

These questions are not our concern in this paper. The studies reviewed here are concerned with global inequality from two points of view. First, it is of interest intrinsically as a measure of the distribution of goods or resources among individuals in the world. Within countries, high levels of inequality are often taken to indicate a lack of fairness in society and governments may act to reduce inequality—for example, through progressive tax-benefit policies. At the global level, there are some redistributive mechanisms, e.g., foreign aid. Moreover, rules governing economic interactions between rich and poor countries, e.g., intellectual property rights over pharmaceuticals, will affect global inequality. A concern for "global justice" will lead to an interest in concept three inequality.

Secondly, changes in global inequality are sometimes portrayed as consequences of "globalization." As trade and financial flows among countries increase, mediated by governments and international institutions that substantially influence the terms of these exchanges, questions of distribution across countries immediately arise. Thus the evolution of global inequality may tell us something about globalization. Although several of the studies reviewed here attribute changes in global inequality to globalization, none presents any causal analysis. Nonetheless, measuring trends in global inequality would be an important preliminary for such an analysis.[5]

Since our main concern is inequality of real income (or consumption) among individuals, it is natural to use PPP exchange rates. However, from a technical point of view, the difference between inequality as measured at PPP rates and at market exchange rates is itself of interest. The divergence reported by studies between trends in inequality at different exchange rates may imply something about economic structure, which we discuss later. Thus we also report global interpersonal inequality at market exchange rates.

We have so far said nothing of concept two inequality, which assigns to each individual in the world the per capita income of his or her country. As the between-country component in the decomposition of concept three inequality, it is useful for explaining the sources of global interpersonal inequality and its changes over time. In addition, some studies (e.g., Firebaugh 1999, 2003) have used it as a downward-biased estimator for concept three inequality.

Finally, it should be noted that our article is concerned with *relative* global inequality and not *absolute* global inequality. Relative inequality remains constant when incomes rise (fall) by the same *proportionate* amount; absolute inequality remains constant when

[5] Any attempt to associate changes in global inequality with globalization in recent decades would, for comparison, have to consider earlier periods in which the world was less "globalized."

incomes rise (fall) by the same *absolute* amount. Anthony B. Atkinson and Andrea Brandolini (2004) compare Bourguignon and Morrisson's (2002) estimates of relative global inequality with indices of inequality that, to varying degrees, measure absolute inequality. Unlike the relative indices, these absolute indices register substantial rises in inequality between 1970 and 1992. This finding supports Martin Ravallion's (2004) argument that disagreements over whether global inequality has gone up or down may partly be due to differing views about the importance of absolute versus relative conceptions of inequality. Absolute conceptions of inequality certainly have some intuitive appeal: for a discussion see Atkinson and Brandolini (2004, pp. 3–4). However, since no other study estimates absolute global inequality, we confine ourselves in this paper to examining relative global inequality.

3. *An Overview of Global Inequality*

As a first cut at estimating international inequality, UNDP (1999) and World Bank (2001) report changes in the ratio of the per capita GDP of the richest countries to that of the poorest countries. This is a measure of concept one inequality as it takes the country as the unit of analysis and per capita GDP (at PPP exchange rates) as the income concept. World Bank (2001, p. 51) reports that in 1960 the per capita GDP of the twenty richest countries was eighteen times that of the twenty poorest countries, while in 1995 the ratio had grown to 37. UNDP (1999, p. 38) notes that the ratio of the per capita GDP of the richest country to that of the poorest country grew from 35 in 1950, to 44 in 1973, and 72 in 1992.[6] Like Pritchett's (1997) well-known analysis, this represents "divergence, big time."

Matters are not so simple, however, when we turn to concept three global interpersonal inequality (see figures 1 and 2 and tables 1 and 2). All studies agree that the *level* is very high: for example, estimates of the Gini coefficient using standard purchasing power parity exchange rates (sourced from the World Bank, Angus Maddison, or the Penn World Tables) in the 1990s lie within the range of 0.63 to 0.686. These levels are comparable to those found within the most unequal countries, such as Lesotho and Namibia, with Ginis of 0.632 and 0.743 respectively, according to World Bank (2007).

In contrast, no consensus emerges concerning the direction of *change* in global inequality in the last twenty to thirty years. For example, Steve Dowrick and Muhammad Akmal (2005) find that the Gini falls from 0.659 in 1980 to 0.636 in 1993 when using standard PPP conversion factors, but that it rises slightly from 0.698 to 0.711 using their own "Afriat" PPP conversion factors (on which more below). Xavier Sala-i-Martín (2006) finds it to decrease from 0.660 in 1980 to 0.637 in 2000, and Bhalla (2002) records a reduction from 0.686 in 1980[7] to 0.651 in 2000. On the other hand, Bourguignon and Morrisson (2002) find no change in the Gini between 1980 and 1992, which remains at 0.657, while their estimate of the

[6] Both Surjit S. Bhalla (2002) and Australian Treasury (2001) object to this procedure, claiming that the measure is biased because it uses different countries in the two years of comparison. They claim that the correct procedure would be to compare the relative incomes of the same groups of countries in the two years and that this results in a decline in inequality. Bhalla (2002, p. 24) states that the income ratio between the richest 20 and poorest 20 countries in 1960 is 23, and that the ratio between these same two groups of countries falls to 9.5 in 2000. This criticism is mistaken. It is an axiom of inequality measures that they are symmetric or "anonymous," i.e., they do not distinguish between individuals (countries) other than by their income level. Inequality measures are functions of the vector of incomes, which are invariant to permutations of the vector, i.e., they are independent of the individual (or country) names attaching to the incomes. The World Bank and UNDP methodology satisfy this axiom, while Bhalla's does not.

[7] This information can be roughly read off the graph in figure 11.1 of Bhalla (2002). This number is also given in table 5.2 on page 80 of the third draft of Bhalla (2002), circulated in December 2001, but the table and this number do not appear in the final published version.

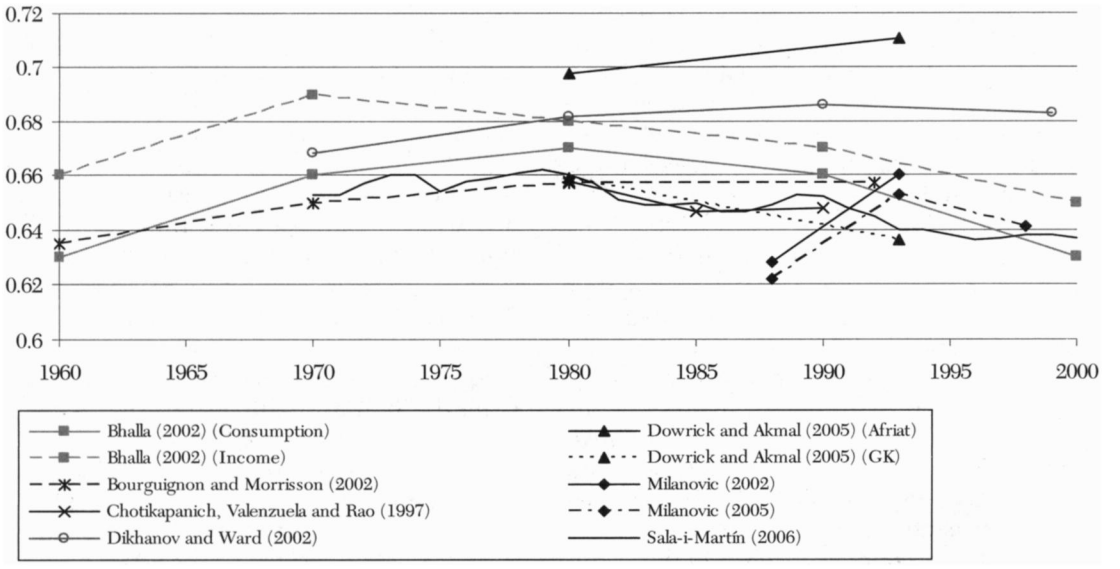

Figure 1. Estimates of Global Interpersonal Inequality at PPP$: Gini Coefficient

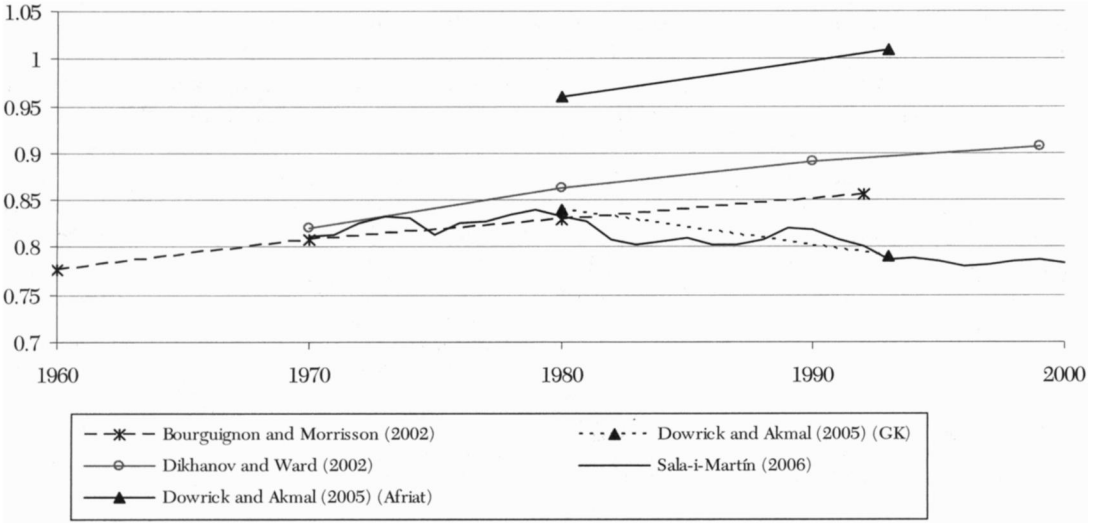

Figure 2. Estimates of Global Interpersonal Inequality at PPP$: Theil *T*

TABLE 1
ESTIMATES OF GLOBAL INTERPERSONAL INEQUALITY: PPP EXCHANGE RATES

	1960	1970	1980	1985	1988	1990	1992	1993	1998	1999	2000
Gini Coefficient											
Bhalla (2002) (Income)[a]	0.66	0.69	0.68			0.67					0.65
Bhalla (2002) (Consumption)	0.63	0.66	0.67			0.66					0.63
Bourguignon and Morrisson (2002)	0.635	0.650	0.657				0.657				
Chotikapanich, Valenzuela and Rao (1997)			0.658	0.647		0.648					
Dikhanov and Ward (2002)		0.668	0.682			0.686				0.683	
Dowrick and Akmal (2005) (GK)			0.659					0.636			
Dowrick and Akmal (2005) (Afriat)			0.698					0.711			
Milanovic (2002)					0.628[b]			0.660[b]			
Milanovic (2005)					0.622[b]			0.653[b]	0.641		
Sala-i-Martín (2006)		0.653	0.660	0.650	0.649	0.652	0.645	0.640	0.638	0.638	0.637
Theil *T*											
Bourguignon and Morrisson (2002)	0.776	0.808	0.829				0.855				
Dikhanov and Ward (2002)		0.821 [74.3%]	0.863 [74.4%]			0.891 [74.2%]				0.907 [70.5%]	
Dowrick and Akmal (2005) (GK)			0.84 [70.9%]					0.79 [70.4%]			
Dowrick and Akmal (2005) (Afriat)			0.96 [71.5%]					1.01 [71.4%]			
Sala-i-Martín (2006)		0.812 [68.6%]	0.833 [68.6%]	0.809 [67.8%]	0.808 [67.8%]	0.818 [68.1%]	0.800 [66.6%]	0.787 [65.8%]	0.785 [64.4%]	0.787 [64.3%]	0.783 [63.8%]
Theil *L* (Mean Log Deviation)											
Chotikapanich, Valenzuela and Rao (1997)			0.855	0.803		0.806					
Dikhanov and Ward (2002)		0.996	1.061			1.021				0.971	
Milanovic (2002)					0.765[b] [75%]			0.873[b] [74%]			
Milanovic (2005)					0.727[b] [72%]			0.817[b] [72%]	0.789 [71%]		
Sala-i-Martín (2006)		0.861 [71.5%]	0.888 [71.1%]	0.847 [68.6%]	0.842 [67.6%]	0.855 [67.5%]	0.833 [65.6%]	0.819 [64.6%]	0.816 [62.0%]	0.819 [61.6%]	0.820 [61.1%]

TABLE 1 (CONTINUED)

	1960	1970	1980	1989	1990	1993	1998	2000
Variance of log-income (Varlog)								
Dowrick and Akmal (2005) (GK)			1.74 [65.3%]			1.51 [62.8%]		
Dowrick and Akmal (2005) (Afriat)			2.21 [70.1%]			2.40 [63.4%]		
Sala-i-Martín (2006)		1.581	1.644	1.593	1.593	1.558	1.585	1.623
Schultz (1998)	1.416 [66.6%]	1.565 [70.7%]	1.524 [71.3%]	1.441 [70.2%]				

Notes: [a] Bhalla specifies numerical estimates only for world income inequality in 1960, 1973, and 2000. However, in figure 11.1 (p. 174) he plots Ginis for world income and consumption inequality for each year during 1950–2000. From this figure we have read off the Gini values to two decimal places for the years reported here.
[b] The estimates for 1988 and 1993 in Milanovic (2002) differ from those in Milanovic (2005) because the common sample is slightly different.
Figures in square brackets show between-country contribution where estimated.

TABLE 2
ESTIMATES OF GLOBAL INTERPERSONAL INEQUALITY: MARKET EXCHANGE RATES

	1965	1980	1988	1992	1993	1998
Gini Coefficient						
Dowrick and Akmal (2005)		0.779 [90.9%]			0.824 [92.4%]	
Korzeniewicz and Moran (1997)	0.749 [91.1%]			0.796 [92.7%]		
Milanovic (2002)			0.782[a]		0.805[a]	
Milanovic (2005)			0.778[a]		0.799[a]	0.794
Theil *T*						
Dowrick and Akmal (2005)		1.25 [77.3%]			1.50 [79.0%]	
Korzeniewicz and Moran (1997)	1.145 [78.8%]			1.321 [85.6%]		
Theil *L* (Mean Log Deviation)						
Milanovic (2005)			1.283 [86%]		1.380 [85%]	1.348 [83%]
Variance of log-income (Varlog)						
Dowrick and Akmal (2005)		3.67 [74.7%]			4.23 [72.2%]	

Notes: [a] The estimates for 1988 and 1993 in Milanovic (2002) differ from those in Milanovic (2005) because the common sample is slightly different.
Figures in square brackets show between-country contribution where estimated.

Theil T index increases from 0.829 to 0.855. Milanovic (2005) finds that the Gini coefficient increases from 0.622 to 0.641 between 1988 and 1998.

Several studies estimate what they call the "Theil" measure of inequality. Unfortunately, the authors are not referring to the same (Theil) inequality index.[8] In Duangkamon Chotikapanich, Rebecca Valenzuela, and D. S. Prasada Rao (1997), Milanovic (2002, 2005), and Yuri Dikhanov and Michael Ward (2002), the "Theil index" refers to the Theil L measure[9,10] or the mean logarithmic deviation (Anand 1983, pp. 89–91), but in Bourguignon and Morrisson (2002), Dowrick and Akmal (2005), Roberto P. Korzeniewicz and Timothy P. Moran (1997), and Sala-i-Martín (2006, 2002a, 2002b), the Theil index refers to the Theil T entropy measure.

Estimated changes in the Theil indices are typically larger than those in the Gini. One explanation might be that changes in the incomes of the richest relative to those of the poorest have been more significant than those in the middle of the distribution. The Gini can be less sensitive to such shifts in the income distribution than either of the Theil indices. The findings in Milanovic (2002) that the ratio of the income of the richest 5 percent to the poorest 5 percent increased from 78 in 1988 to 114 in 1993, and in Bourguignon and Morrisson (2002) of an increase in the global income share of the top 5 percent between 1980 and 1992, are consistent with this explanation.

Four of the studies also calculate global inequality at market exchange rates. The level found is, not surprisingly, substantially higher than when PPP incomes are used and all four studies also report an increase over time. Dowrick and Akmal find that the Gini rises from 0.779 to 0.824 between 1980 and 1993, Milanovic (2002) from 0.782 to 0.805 between 1988 and 1993, Milanovic (2005) from 0.778 to 0.794 between 1988 and 1998, and Korzeniewicz and Moran from 0.749 to 0.796 between 1965 and 1992.

With increasing globalization, one would expect market exchange rates to move closer to PPP exchange rates (as countries trade larger proportions of their GDP). The apparent divergence over time between inequality measured at market and at PPP exchange rates thus requires some explanation. Dowrick and Akmal (2005) attempt to address this question but in our view, discussed below in section 4.2, the relationship between changes in global inequality at PPP and at market exchange rates merits further research.

4. *Methods and Data*

The wide range of findings across the different studies requires explanation. In this section, we discuss three issues that bear on all the studies and that correspond to dimensions in which the studies differ: the use of national accounts versus household surveys to estimate national mean income or consumption; the choice of exchange rate in constructing a global distribution from national distributions; and the definition of the global distribution. Questions that are specific to methods used by individual studies are dealt with in section 5, which reviews the studies, unless they serve to clarify general issues.

4.1. *Scaling Within-Country Distributions*

Just as national inequality is estimated using nationally representative household

[8] The Theil T index is an income-share weighted average of the logarithmic difference between each person's income and mean income. The Theil L measure is the simple average of the logarithmic difference between mean income and each person's income; hence Theil L is also referred to as the mean logarithmic deviation (MLD). For formulas and discussion of the two Theil indices and the other inequality measures in table 1, see Sudhir Anand (1983, pp. 303–16).

[9] In discussing the "Theil" index, Dikhanov and Ward provide the formula for the Theil L measure. They also estimate what they call the "Theil 2" index but do not provide any formula for it. We assume it refers to the Theil T index.

[10] Milanovic (2002) does not specify which Theil index he uses, so we contacted the author directly for this information.

surveys, so the measurement of global inequality would ideally be based on a globally representative household survey. In practice, global inequality is estimated by aggregating national surveys. Typically, the papers that estimate global inequality do not use primary data from household surveys, but rather compilations of secondary data. Most have used the compilation by Klaus Deininger and Lyn Squire (1996), which reports only the Gini coefficient and quintile shares of national distributions. These data on within-country inequality are then combined with national accounts data on countries' mean GDP (or consumption) per capita. Thus, the relative distribution in each country is based on (secondary) data from household surveys, which is scaled to a national accounts mean. The only author who does not scale to national accounts means is Milanovic. Unlike all other studies, Milanovic (2002, 2005) uses primary data on income (or consumption expenditures) from household surveys directly—without scaling them—to construct his world distribution of income.[11]

The lack of estimates of other summary statistics, including survey means, in Deininger and Squire (1996) is not a reason to scale to national accounts means. Researchers could obviously have returned to the sources, cited in Deininger and Squire, to obtain the means. The practice of using national accounts means has to be justified on substantive grounds. In this subsection, we discuss the issues concerning the use of national accounts versus survey means in estimating global inequality.

National accounts (NA) data are typically available on an annual basis and go back much further than household surveys. But the use of NA data in the estimation of global inequality raises two questions. The first is whether GDP is the appropriate NA category to scale up to. The second is the more fundamental question of whether NA estimates of mean income (or consumption expenditure) are preferable to estimates obtained directly from the surveys.

In the national accounts, GDP at market prices is defined from the expenditure side as final consumption expenditure, plus gross capital formation,[12] plus exports minus imports (System of National Accounts (SNA) 1993, p. 155). Final consumption expenditure is the aggregate of consumption expenditures of households, "nonprofit institutions serving households"[13] (NPISHs), and government (SNA 1993, p. 353). In addition to GDP, the other option available for scaling within-country distributions is the sum of household final consumption expenditure (HFCE) and final consumption expenditures of NPISHs. For most countries, we cannot disaggregate these two categories of final consumption expenditure.[14] For convenience, we will refer to the aggregate simply as HFCE.

There are two possible alternative categories to GDP and HFCE defined in the 1993 System of National Accounts, but most countries do not report them. The first is Household Disposable Income (SNA 1993, p. 186), which is household incomes (including production for own consumption) after taxes and net cash transfers (without transfers in kind). The second is Household Actual Final

[11] This parallels the World Bank's method for calculating poverty (Shaohua Chen and Ravallion 2001, 2004).

[12] Gross capital formation is defined as gross fixed capital formation, plus changes in inventories, plus acquisitions less disposals of valuables (such as precious stones and metals, and works of art, acquired as "stores of value" and not for production or consumption) (SNA 1993, p. 353).

[13] NPISHs are defined as nonprofit institutions "which provide goods or services to households free or at prices that are not economically significant." These include associations of persons providing goods or services for the benefit of the members themselves (such as trade unions, political parties, and religious societies), and charities, relief, or aid agencies that are created for philanthropic purposes (SNA 1993, pp. 95–96).

[14] Angus Deaton (2005) reports that "in the United Kingdom, NPISH in 2001 was 3.9 percent of total consumption, almost double the 1970 share of 2.1 percent" (p. 15). He adds that the share may be higher in poorer countries. However, data on the subject do not appear to be available.

Consumption (SNA 1993, p. 216), which is defined as goods and services that are acquired by households, either through their own expenditures or through social transfers in kind (e.g., health and education services paid for by government or charities).[15] The lack of data on these two categories for most countries implies that they are not an option for scaling survey means in the estimation of global inequality. In contrast, like GDP, the category of HFCE (including the final consumption expenditures of NPISHs) is reported by almost all countries and is published in the IMF's *International Financial Statistics*.

Given the options of scaling household surveys to GDP or HFCE, what is to choose between them? If we are interested in aggregate household consumption expenditure then HFCE is obviously a better measure to use than GDP.[16] But we may be as interested in household income as in household consumption, and if this is the case then we have to ask whether GDP is an appropriate measure. In our view, it is not a suitable measure of household income. GDP includes depreciation, retained earnings of corporations, and the part of government revenue (taxes) that is not distributed back to households as cash transfers. For illustration, we can take the example of the United States, which is one of the few countries that does report measures of household income (referred to as personal income) in its National Income and Product Accounts (NIPA).[17] In 2006, U.S. GDP in billions was $13,246.6, personal income $10,891.2, and disposable personal income $9,529.1 (NIPA tables 1.1.5 and 2.1). Disposable personal income was therefore only 72 percent of GDP. While most other countries do not provide national accounts data on disposable household income, Deaton (2005, p. 4) compares household income as measured by surveys with GDP. He reports that, on average across 272 surveys, household income amounts to only 0.57 of GDP, with a standard deviation of 0.20 (with the population-weighted average at 0.54 of GDP).

Indeed, aggregate disposable household income may be better approximated even by HFCE than by GDP. Deaton (2005) notes that "much of household saving may not be done by households, but by corporations, government, or foreigners, so that household income may be closer to household consumption than to national income" (p. 4). Across 266 household surveys he reports that on average household income amounts to 90 percent (101 percent population-weighted) of HFCE. This closeness of fit may, of course, be due to different biases—e.g., arising from incomes being underestimated in surveys and household savings being positive—approximately cancelling each other out.

In conclusion, we can see no justification for scaling household income (or consumption) from surveys to GDP. HFCE is the only other widely available national accounts category. By including expenditures of NPISHs, HFCE contains some systematic upward bias as a measure of household consumption. The magnitude of this bias is not known for most

[15] Bettina Aten and Alan Heston (2004) observe that "The OECD countries have all adopted the 1993 SNA, and PWT 6.1 (2002) but most developing countries do not yet provide Household Actual Final Consumption" (p. 6).

[16] If we wanted to include government transfers in kind (e.g., health and education services) then household actual final consumption would be the appropriate category. Since data on this category are generally not available, it might be argued that GDP can be used to approximate it. However, since GDP includes gross capital formation, a better measure would be total final consumption expenditure (of households, NPISHs, and government). But scaling survey distributions to this category implies that the benefits of government expenditure are distributed in proportion to household income or consumption expenditures, which is an assumption without basis. None of the papers reviewed here use this category.

[17] In NIPA, the category of "personal income" is defined as "the income received by persons from all sources—that is, from participation in production and from current transfer receipts from both government and business," while "disposable personal income" is "personal income less personal current taxes." It is "the income available to persons for spending or saving" (Bureau of Economic Analysis 2007, p. 5). This disposable personal income corresponds to Household Disposable Income in SNA 1993.

countries, but is unlikely to be large as the upward bias of GDP as a measure of household income. Dikhanov and Ward (2002) scale survey distributions to what they refer to as "personal consumption expenditure" (p. 14), while Bhalla (2002) scales to "private final consumption expenditure" (p. 217),[18] both of which we take to mean HFCE. The question still remains, of course, whether scaling to HFCE is preferable to using survey means directly.

The choice between survey means and per capita HFCE can be important in measuring global inequality because the ratio of survey to NA household consumption varies across countries and over time. Deaton (2005, p. 4) reports that, on average across 277 surveys, household consumption in surveys is 0.86 of HFCE, with a standard deviation of 0.31 (with a population-weighted average of 0.78). This will have implications for the measurement of global inequality: for example, Milanovic (2005, p. 118) finds that if he scales survey income or consumption means to GDP per capita rather than using the survey means themselves, then his 1988 estimate rises by nearly two Gini points.

Survey household expenditure differs from the NA category of HFCE in both concept and method of estimation. In terms of concept, HFCE includes imputed values of financial intermediation services and consumption by NPISHs. HFCE also includes imputed rents from owner-occupied housing, which is rarely estimated in household surveys. Both include household production for own consumption,[19] but it should be noted that neither survey expenditure nor HFCE includes imputed values of government-provided services such as healthcare and education.[20]

The two categories differ radically in their method of estimation.[21] To calculate HFCE, the NA typically starts with an estimate of national production of a commodity such as rice from crop-cutting data, aerial or farm surveys, etc. As such surveys are conducted infrequently, gross production figures may have to be estimated without up-to-date information. Moreover, the methods used to arrive at these figures are not applied uniformly and can be unreliable. From an estimate of national production thus generated, government consumption and firms' consumption are subtracted. The residual is attributed to households. Data on government consumption may be adequate, but firms' consumption is typically poorly estimated. It is often based on outdated firm surveys and extrapolations or assumed changes over time. In India, survey and NA mean expenditure have diverged (Deaton 2005, p. 8) and this is partly due to the underestimation by NA of firms' consumption of intermediate goods. This has led to double-counting where, for instance, the edible oil consumed in restaurant meals was attributed to HFCE under both the "edible oil" category and the "restaurant meals" category.[22]

NA estimates of HFCE are thus indirect and subject to three sources of error: the initial estimate of aggregate production,

[18] In other places (e.g., p. 128), Bhalla refers ambiguously to "NA consumption." Note that he separately estimates global income inequality by scaling survey distributions to GDP per capita (pp. 103–04, footnote 1).

[19] Both also include goods and services that are bartered. In surveys they count as both income and consumption (the challenge being to impute an appropriate value for the goods, as in production for own consumption), while in national accounts the goods and services will have been included in the estimate of total production.

[20] Hence household consumption expenditure from surveys is conceptually closer to HFCE than to household actual final consumption. Since household actual final consumption includes household consumption paid for by government and NPISHs, it is "a better indicator of [household] living standards than their final expenditure alone" (SNA 1993, p. 205). As explained above, most countries do not report it in their national accounts. It should also be noted that governmental provision of benefits in kind to households is typically greater in richer countries, so that the use of survey incomes (consumption) could lead to an understatement of "full" income (consumption) differences—including such benefits in kind—between richer and poorer countries.

[21] Much of this paragraph closely follows Deaton (2003, pp. 367–68).

[22] See A. C. Kulshreshtha and Aloke Kar (2005).

the estimate of government consumption, and the estimate of firms' consumption. There is no reason to suppose that the data and methods used to estimate these, which include surveys of various kinds, are more reliable than household surveys. Moreover, their sources and methods are generally less well-documented (in terms of the surveys used, how and when they were conducted, etc.) than household surveys. Finally, as it is defined a residual, the errors in the estimate of HFCE will tend to get compounded.

Household surveys measure personal income or expenditure directly. Two major problems with household surveys are that the rich disproportionately fail to respond and, when they do respond, they tend to underreport their income and expenditure. On the other hand, the very poor and marginalized, particularly the homeless or those living in remote rural areas, tend to be excluded from the sample frame and are thus likely to be underrepresented. In most countries, the net result is that mean income or expenditure in surveys is lower than per capita HFCE in NA (Deaton 2005). In India there has been heated debate on the size and source of the divergence between survey and NA means in the context of poverty estimation (e.g., Bhalla 2002, Deaton 2005, Deaton and Valerie Kozel 2005, Ravallion 2000). One factor explaining the divergence is that when within-country inequality rises, and with it the income share of the rich, as has occurred in both India and China, undersampling of and underreporting by the rich implies a growing underestimation of average household income and expenditure.

Underreporting by the rich will also lead to a downward bias in measured within-country inequality, although the effect of undersampling is theoretically ambiguous (Anand 1983, pp. 343–44).[23] The impact of underestimating mean income or expenditure in countries will depend on how the degree of underestimation varies with the level of the actual mean, which will determine the direction and magnitude of the bias in between-country inequality. We are not aware of any attempts to estimate the bias in measured global inequality due to undersampling and underreporting of incomes in household surveys.

Both methods of estimating global inequality—taking incomes or expenditures directly from surveys, or using NA means and within-country distributions from surveys—suffer from the same underestimation of within-country inequality because they both use (the same) household surveys for their estimates of within-country distributions. It is between-country inequality that is affected by the choice of method. If the use of NA means entails scaling-up survey means proportionately more (less) for poorer than for richer countries, then between-country inequality based on NA means will be lower (higher) than that based on survey means.[24]

It is clear that there are estimation errors in both sources of data. We do not know their relative magnitude, and in particular there is little reason to believe that NA are more accurate than surveys in measuring household consumption. Given that we take within-country distributions from surveys, it seems anomalous that we should seek an alternative source for the *means* of these distributions.

To address both the undersampling and underreporting problems, a possible route may be to estimate parametrically within-country distributions from the unit-record information contained in each household

[23] Anton Korinek, Johan A. Mistiaen, and Ravallion (2006) find that selective nonresponse in the 2004 U.S. Current Population Survey biases the measured Gini downwards. From their finding that "the average state-level income is negatively correlated with response" (p. 42), they estimate a negative relationship between household income and probability of response. Correcting for nonresponse raises the Gini from 0.448 to between 0.492 and 0.498, depending on the specification of their model (pp. 42, 45).

[24] As noted earlier, Milanovic (2005) finds that scaling up survey means to GDP per capita increases measured between-country inequality in 1988. We are not aware of any attempts to compare global inequality estimated using survey means with that estimated using HFCE.

survey. For example, one could specify a distribution for each country that incorporates a plausible upper tail and estimate it from the household survey data. The estimated distribution would then provide us with corrected estimates both for average income and for the level of inequality. This would appear to be superior to the scaling-up procedure that applies the same multiplicative factor to adjust all incomes in the survey.

The choice between survey and NA mean, and that between HFCE and GDP, has not been adequately addressed in the literature on global inequality. Bhalla (2002) discusses some of the issues and argues that in the case of India HFCE is more accurate than consumption expenditure from household surveys. Deaton (2003), K. Sundaram and Suresh D. Tendulkar (2003), and Ravallion (2000) disagree with this conclusion. Milanovic (2002) discusses the choice between survey mean and GDP per capita and, as stated earlier, reports the difference that this choice makes to estimated inequality (Milanovic 2005). Sala-i-Martín (2002b), a precursor to Sala-i-Martín (2006), briefly discusses the choice between income and consumption in the measurement of poverty but not in the measurement of inequality; in Sala-i-Martín (2006, p. 357, footnote 5) he makes a reference to Deaton (2005) on the subject. For his estimates of global income inequality Sala-i-Martín (2006) resorts to GDP per capita.

In our view, then, NA means are not in general more reliable than household survey means. Moreover, GDP does not correspond even in principle to household income, while the residual nature of estimates of HFCE makes them suspect. There is also a basic incongruity in assuming that the relative within-country distributions are measured acceptably well by surveys but their means are not. Survey means directly measure the average level of the variable of interest, even if imperfectly, and are therefore the appropriate complement to survey distributions. When household surveys are not available, as is the case going further back in time, we believe that researchers need to be more circumspect regarding the accuracy of their estimates,[25] an issue to which we return in detail in section 7 below.

4.2. *PPP Exchange Rates*

The estimation of global inequality requires the conversion of national currencies into a common numeraire so that we can aggregate national distributions into a global distribution. One possibility is to use market exchange rates, but these suffer from the well-known "traded-sector bias": market exchange rates are affected by the prices of traded goods across countries but do not reflect domestic prices of nontraded goods. Since the relative price of nontraded to traded goods tends to be lower in poorer than in richer countries, valuing domestic incomes at market exchange rates will undervalue incomes in poorer relative to richer countries and impart an upward bias in measured inequality between countries. For this reason the preferred method is to use PPP exchange rates, which are supposed to reflect relative purchasing power across countries better than do market exchange rates. But few users have a good idea of how PPP rates are constructed or their interpretation. There are two commonly used methods for constructing PPP rates: Geary-Khamis (GK) and Eltetö-Köves-Szulc (EKS). The GK method is used by the Penn World Tables (PWT) and Maddison (1995, 2001), and was formerly used by the World Bank, while the EKS method has been used by the World Bank for its more recent estimates of PPP incomes. A third method of constructing PPPs, designated "Afriat," was introduced by Dowrick and John Quiggin (1997) and has been used to measure global inequality by Dowrick and Akmal (2005).

[25] Indeed, in using NA estimates of mean incomes to calculate global inequality further back in time, household surveys are still required for within-country distributions. If appropriate distributional data from the past are not available, then we have to admit that we just do not have the information to estimate global inequality.

The GK method consists in estimating an "international price vector" for commodities at which the vector of outputs of each country is valued to yield its "real GDP." The method consists of simultaneously solving two sets of equations: the first set determines the PPP exchange rates for each country (with the United States typically chosen as numeraire), and the second set determines the "international price" of each commodity. A country's PPP exchange rate is defined as the value of its output at domestic prices divided by the value of its output at "international prices." The "international price" of a commodity is defined as a weighted average across countries of its domestic price in each country divided by the country's PPP exchange rate, with the weights being the country's share of output of the commodity in total global output. Hence, for $i = 1,\ldots,m$ commodities and $j = 1,\ldots,n$ countries, with p_{ij} and q_{ij} the domestic price and output of commodity i in country j, respectively, the PPP exchange rate PPP^j_{GK} of country j, and the international price π_i of commodity i, are defined by:

$$PPP^j_{GK} = \frac{\sum_{i=1}^{m} p_{ij}q_{ij}}{\sum_{i=1}^{m} \pi_i q_{ij}} \qquad j = 1,\ldots,n;$$

$$\pi_i = \sum_{j=1}^{n} \frac{p_{ij}}{PPP^j_{GK}} \left[\frac{q_{ij}}{\sum_{j=1}^{n} q_{ij}} \right] \qquad i = 1,\ldots,m.$$

This system of $n+m$ equations is solved for the $n+m$ unknowns in a notional international currency. To denominate international prices and PPP exchange rates in U.S. dollars, with the United States being country n, one can substitute the equation $PPP^n_{GK} = 1$ for the n^{th} PPP_{GK} equation. While U.S. dollars are thereby set as numeraire, relative prices remain relative "international prices" and not relative U.S. prices.

The EKS method estimates a PPP exchange rate by generalizing the Fisher index between countries and does not involve the construction of a vector of "international prices." Recall that the Laspeyres and Paasche indices for country A's prices relative to B's prices are constructed by dividing country A's prices by country B's prices, in the case of Laspeyres weighting both sets of prices by country B's output, and in the case of Paasche weighting both at country A's output. The Fisher index is then the geometric mean of these two indices. As it stands, the Fisher index is not suitable for multilateral comparisons of prices between more than two countries because the bilateral comparisons are not transitive: that is, the price level of India relative to the United States will not be equal to the price level of India relative to Brazil times the price level of Brazil relative to the United States. The EKS method generalizes the Fisher index to create a transitive set of real exchange rates simply by averaging all of the Fisher indices for a country. Let F^{kj} be the Fisher index of the prices of country k relative to those of country j. Then the PPP exchange rate for country j is

$$\ln PPP^j_{EKS} = \frac{1}{n}\sum_{k=1}^{n} \ln F^{kj}.$$

Since $\ln F^{kj} = -\ln F^{jk}$ we have

$$\ln PPP^j_{EKS} - \ln PPP^l_{EKS} = \frac{1}{n}\sum_{k=1}^{n}(\ln F^{lk} + \ln F^{kj})$$

from which it follows that the exchange rates are transitive. This procedure assigns to each country a real income number that allows relative income comparisons between countries. These numbers are then scaled so that U.S. output is equal to its actual US$ value,

with other countries' outputs then denominated in dollars.

The GK and EKS methodologies are evidently very different.[26] Moreover, there is reason to believe that the incomes of poorer countries relative to richer countries will be biased upwards using GK owing to substitution bias (also known as Gerschenkron bias). This refers to the fact that valuing the output of both country A and country B at country B's prices will lead to an overestimation of the income of country A relative to that of country B. The GK method constructs "international prices" by weighting a country's price by the country's share in world output, which leads to relative "international prices" being closer to those obtaining in richer than in poorer countries. Using the GK index with data in PWT 5, Nuxoll (1994) finds that "income indexes based on international prices closely resemble indexes based on the prices of some moderately prosperous country. The closest fit is Hungary; the second closest is Yugoslavia" (p. 1431). Dowrick and Akmal (2005) claim that this results in the incomes of poor countries being overestimated by more than the incomes of rich countries, and that global inequality will therefore be biased downwards. Robert Ackland, Dowrick, and Benoit Freyens (2004) find that the GK method overvalues the incomes of poorer countries compared to EKS: if log per capita GDP from GK is regressed on log per capita GDP from EKS then the slope is 0.94 and is significantly less than 1.

Although EKS is not subject to the same substitution bias as GK, Dowrick and Akmal (2005) use a third method for constructing PPP incomes that is also not subject to this bias—the Afriat PPP (Dowrick and Quiggin 1997, Dowrick 2002). The papers by Dowrick et al. make much use of Sydney N. Afriat's (1981) theorem that the existence of a set of Afriat PPPs is equivalent to the existence of a representative consumer with a common homothetic utility function, which rationalizes all of the observed consumption baskets across countries. While this may appear to be a satisfying justification for Afriat PPPs, the problem, as Ackland, Dowrick, and Freyens (2004, p. 18) find, is that, in the 1996 ICP, only 80 of 115 countries can be aggregated into a set that does not violate the hypothesis of common homothetic preferences. No set of Afriat PPPs exists for all 115 countries and, as Ackland et al. point out, "the fact that nearly one third of the ICP countries do not satisfy the test is a major weakness in applying the Afriat approach to constructing a comprehensive multilateral index" (p. 18). We would add that it is unlikely that any country has homothetic preferences—certainly no country that satisfies Engel's Law. The inapplicability of the Afriat approach to a third of countries in the ICP significantly undermines its theoretical advantages.

Dowrick and Akmal (2005) estimate global inequality in 1980 and 1992 using GK (from PWT), Afriat, and market exchange rates. They find that GK inequality declines, Afriat inequality rises slightly, and market exchange rate inequality rises substantially (see our table 1). They argue that the divergence over time of GK from Afriat inequality, and of market exchange rate from Afriat inequality, are due to increases over time in, respectively, substitution bias and traded-sector bias. These biases have grown, they contend, because price structures have diverged over time. They estimate trends in price similarity across countries, and find that price structures became less similar over the period 1980 to 1991 (Dowrick and Akmal 2005, p. 213, figure 5); hence they argue this would cause both traded-sector and substitution bias to increase. As Afriat PPPs do not

[26] However, from the above formulas it can be seen that both methods are subject to the following complication. For any bilateral comparison of PPP income or consumption between two countries A and B, a change in output or prices in a third country C will affect the bilateral comparison between A and B. Prices and output in all countries affect the PPP of each country, through the international price vector in GK, and directly in the case of EKS.

suffer from either traded-sector or substitution bias, the rise in market exchange rate inequality and the fall in GK (PWT) inequality could both in principle be explained by increases in the two biases.

We have reservations about the argument regarding traded-sector bias. The claim that this bias increases as "price structures" become less similar assumes that the price structure in question is the relative price of traded to nontraded goods. However, the price vector that Dowrick and Akmal (2005, p. 212) use to measure price divergence across countries comprises the relative prices of private consumption, investment, and government consumption. These three categories each comprise both traded and nontraded goods, so Dowrick and Akmal's empirical exercise does not establish divergence in relative prices of traded to nontraded goods across countries. With globalization, moreover, we would expect the traded sector to expand relative to the nontraded sector, which would contribute to a decline in the bias. The finding of a rise in global inequality at market exchange rates would thus seem to require further explanation.

The argument regarding rising substitution bias does seem valid, however, and is consistent with the fact that Dikhanov and Ward (2002), who use EKS PPPs uniformly, find a greater increase in inequality according to most measures than the studies using GK PPPs—viz., Chotikapanich, Valenzuela, and Rao (1997), Dowrick and Akmal (2005), Schultz (1998), Bourguignon and Morrisson (2002), and Sala-i-Martín (2006) (see our table 1). Their estimates are also higher in almost all years and according to all indices than the estimates based on GK PPPs.

Although any study of global inequality should use a set of consistent PPP exchange rates, two of the studies combine PPP rates estimated by different methods. Bhalla (2002) appears to have used both World Bank (EKS) PPPs and PWT (GK) PPPs, although his book is unclear about his sources (see section 5.4). Milanovic (2002) apparently also uses both PWT and World Bank PPP data.[27] It is not clear what effect this mixing of PPPs based on different methodologies will have on estimated global inequality, but the PPPs from different sources are simply not comparable.[28]

It should be apparent from our discussion that the construction and use of PPPs is more complicated than many researchers acknowledge. EKS does not suffer from the problems faced by GK or Afriat discussed above and, given the alternatives available, seems a more appropriate exchange rate to use in the estimation of global inequality. Dikhanov and Ward (2002) is the only study to use EKS PPPs exclusively. More generally, inadequate recognition of the basis of different PPP exchange rates and their inappropriate mixing will add another layer of doubt to estimates of global inequality. We return to PPPs in section 7 below where empirical problems in their calculation, which are distinct from these conceptual issues, are seen to lead to further uncertainty in the estimates.

4.3. *What Distribution?*

Most studies refer to the "global income distribution," but ambiguity remains regarding the distribution that is being estimated. Any distribution must be defined with respect to a given income concept and population unit, but the household surveys used to construct global distributions are a mixture of distributions of income and consumption, defined

[27] Milanovic (2002, p. 62, note to table 6) cites sources just for four countries, the sources being both PWT 5 and the World Bank (cited as "ICP tables provided by Yonas Biru (World Bank)") for his 1988 estimates. Milanovic (2005) cites only EKS PPPs (p. 105, note 88). However, inspection of his (2002) and (2005) datasets shows that the same PPP conversion rates were used for his 1988 estimates in both cases. The (2005) dataset is downloadable from his World Bank website. We downloaded the (2002) dataset earlier, but it no longer appears to be on the website.

[28] For instance, the PWT GK estimate of India's GDP per capita in 1990 is some 15 percent larger than the World Bank's EKS estimate (reported in World Development Indicators Online, June 2007).

with respect to individuals and households. It is therefore not clear exactly what type of global distribution emerges from combining these surveys. Concept three global interpersonal inequality takes the individual as the population unit, but the Deininger and Squire dataset mixes distributions of households and individuals.[29] Moreover, the income concept is not specified in several studies. In this sense, their global distributions are not well-defined and it is unclear whether their final estimates are of consumption or income inequality.[30] This is important because consumption inequality can move in a different direction from income inequality.

To expand their datasets, the studies mix different types of within-country survey distributions, with the result that the data used are not comparable in various respects. Some distributions are of income, others of consumption expenditure. In some country surveys, incomes are gross-of-tax and in others net-of-tax; for some they refer to cash incomes and for others certain items of income-in-kind are included. The rental value of owner-occupied housing is imputed in some surveys but not in others. The population unit of the distribution can be individuals or households (sometimes families), and these units may be ranked in a variety of ways—for example, individuals ranked by income received, individuals ranked by household income per capita (or per equivalent adult), households ranked by household income per capita (or per equivalent adult), households ranked by total household income. The population unit and ranking concept used to construct the distribution can make a huge difference to measured inequality. For example, Anand (1983) found that the income share of the lowest 40 percent varied from 9.6 percent to 17.7 percent for differently defined distributions of income from the *same* Malaysian household survey.

Atkinson and Brandolini (2001) discuss some of these issues in their review of "secondary" datasets used in studies of income inequality. On the basis of a detailed analysis of distribution data for OECD countries, they find that problems of comparability, including those described above, are present even in the "high quality" subset of the Deininger and Squire (1996) compilation. Atkinson and Brandolini (2001) conclude that: "users could be seriously misled if they simply download the Deininger and Squire 'accept' series [i.e., the 'high quality' subset]. Moreover, if the user goes on to utilize the variable in econometric work, then it may make a significant difference to empirical findings" (pp. 777–78). In addition to comparability problems across countries, Atkinson and Brandolini find that changes in survey definitions over time within a given country "may affect not just the *level* but also the *trend* in inequality" (p. 780). The Deininger and Squire (1996) dataset is used by Bhalla (2002), Chotikapanich, Valenzuela, and Rao (1997), Dowrick and Akmal (2005), Sala-i-Martín (2006, 2002a, 2002b), and Schultz (1998). Similar issues arise in respect of the other studies, whose distribution datasets are also subject to noncomparabilities. Recent World Bank distributional data, described in Chen and Ravallion (2004), are subject to fewer problems of noncomparability than Deininger and Squire (1996)—for instance, all country distributions have the individual as the unit—but other problems remain, such as the unavoidable mixing of income and consumption distributions.

[29] The proportion of surveys in Deininger and Squire that take the household as the unit rather than the individual declines over the 1990s, which may imply a bias in estimated changes in inequality. We thank an anonymous referee for pointing this out.

[30] Bhalla (2002) is an exception in that he estimates global income and consumption distributions separately. Schultz (1998) also includes a dummy variable in his regression to estimate within-country inequality to indicate whether a survey is an income or consumption survey. However, Atkinson and Brandolini (2001), on the basis of their analysis, "doubt whether a simple additional or multiplicative adjustment is a satisfactory solution to the heterogeneity of the available statistics" (p. 790).

TABLE 3
DATA SOURCES FOR ESTIMATES OF GLOBAL INTERPERSONAL INEQUALITY AT PPP$

	Within-country inequality data	PPP source
Bhalla (2002)	Own dataset	GDP PPPs from *World Development Indicators* and PWT 5.6
Chotikapanich, Valenzuela and Rao (1997)	Deininger and Squire (1996)	PWT 5.6
Dikhanov and Ward (2002)	Milanovic (2002)	Consumption PPPs from World Bank
Bourguignon and Morrisson (2002)	Own dataset	GDP PPPs from Maddison (1995)
Dowrick and Akmal (2005) (PWT PPPs)	Deininger and Squire (1996)	GDP PPPs from PWT 5.6
Dowrick and Akmal (2005) (Afriat PPPs)	Deininger and Squire (1996)	Own calculations of Afriat index for GDP PPPs
Milanovic (2002, 2005)	Own dataset	Consumption PPPs from PWT and World Bank
Schultz (1998)	Deininger and Squire (1996)	GDP PPPs from PWT 5.5
Sala-i-Martín (2006)	Deininger and Squire (1996) and UNU–WIDER	GDP PPPs from PWT 6.0

5. *The Studies*

We now turn to a detailed examination of the individual papers. Having reviewed the methodological issues that are common to all the studies, in this section we discuss specific methods used by the authors to calculate global inequality. Their individual estimates of global inequality are reported in our tables 1 and 2 above. A summary of their data sources is provided in our table 3.

5.1. *Milanovic (2002, 2005)*

Milanovic (2002) uses income or expenditure taken directly from national household surveys to construct a "true," in the sense of directly observed, world distribution of income/consumption. Using PPP exchange rates for consumption, from both the Penn World Tables (PWT) and the World Bank, he finds that inequality increases over the five-year period 1988–93 (see our table 1). Moreover, the 1988 distribution Lorenz-dominates the 1993 distribution and, hence, will show less inequality for all measures in the Lorenz class of indices (Anand 1983, pp. 339–40). His later book, Milanovic (2005), updates the estimates to 1998. He finds that inequality falls over 1993–98, but remains some two Gini points higher in 1998 than in 1988 (see our tables 1 and 2).

Milanovic's (2002) data comprise a total of 216 country surveys benchmarked to the two years 1988 and 1993, which are obtained from the World Bank and other sources.[31] Milanovic's (2005) study, which includes the benchmark year 1998, has a sample of 345 country surveys. Unlike every other study,

[31] The sample that is common to both 1988 and 1993 consists of 91 countries; in addition, for 1988 he has data for another 10 countries, and for 1993 for another 28 countries. Thus the total number of countries in 1988 is 101 and in 1993 it is 119. The total of 220 country-years is larger than the 216 country surveys. This may be due to his splitting of four large countries (China, India, Bangladesh, and Indonesia) into urban and rural areas in both years and treating them as different countries. He similarly splits Pakistan in 1988 but not in 1993. However, even when we count these observations as distinct surveys for different country-years we are unable to make the numbers tally.

Milanovic constructs inequality estimates over time for a *common sample* of countries. The common sample is slightly different in the two studies but in both cases covers about 84 percent of the world's population. Most of his within-country distributions are described by at least ten quantile shares or income groups, and he assumes that each individual or household within a group has the same income.

Milanovic's estimates raise a number of questions. First, countries in the common sample have different numbers of income groups in the different benchmark years, with the average number of data points (income groups) per country-year standing at 10.8 in 1988, 11.4 in 1993, and 15.1 in 1998. Hence, the underestimation due to the assumption of equal incomes within income groups would be expected to be different in each year. Secondly, the measured distribution within China is of concern. Milanovic (2002) has several income groups in rural China each containing more than 100 million people, with the largest containing 180 million in 1993 (and 175 million in 1990 for the benchmark year 1988).[32] The presence of such large groups could lead to possibly substantial downward biases in measured inequality.

Thirdly, to achieve finer-grained distributions, Milanovic (2002, p. 60) states that he splits four large countries in 1993 and five in 1988, including China and India, into rural and urban areas (for which he has separate distributions) and treats these observations as different "countries."[33] However, in two of these countries—Bangladesh[34] and Indonesia—the corresponding urban and rural income groups have near-identical mean incomes for all but the top and bottom income groups (presumably because the same absolute income intervals were used to code urban and rural incomes). Hence, even though the urban and rural population shares in each income group are different, the urban–rural disaggregation adds almost no information.[35] Note that this is not the case for India and China, where it would really matter: their income groups have different mean incomes in rural and urban areas.

Milanovic (2002, p. 61) also reports using different PPP rates for rural and urban China to take account of price differences between the strata: "For China, in 1993, I use the rate reported in the International Comparison Programme (ICP) for urban areas only (since the rate itself was obtained from surveys conducted in two cities: Guangdong and Shanghai), and reduce the price level in rural areas by an estimated 20 percent (see Yao and Zhu 1998, p. 138)." Shujie Yao and Liwei Zhu (1998) themselves suggest adjusting rural incomes upwards relative to urban incomes by "15 per cent for low cost of living in the countryside" (p. 138). This is equivalent to adjusting rural prices relative to urban prices down by 13 percent, not 20 percent.

5.2. *Bourguignon and Morrisson (2002)*

Bourguignon and Morrisson (2002) estimate global inequality back to the nineteenth century, starting in 1820 and ending in 1992. They assemble data for thirty-three countries

[32] This largest group is equal in size to the combined populations of the fifty smallest countries in his dataset, which between them have more than five hundred income groups.

[33] This implies that his "between-country" component of global inequality actually includes some within-country inequality.

[34] In his on-line (2002) dataset, we could find Bangladesh split into urban and rural areas only in 1988, not 1993.

[35] In his (2002) dataset, the income distribution for Egypt in 1988 is also shown separately by urban and rural areas, with five income groups in each sector. The income groups are shown as quantiles (bottom two quintiles, middle 40 percent, and top two deciles) for each sector. Surprisingly, the second-to-top decile for the urban and rural sectors have identical mean incomes, and for all quantiles other than the bottom quintile the urban and rural mean incomes are very close. Effectively this means he only has five or six income groups for Egypt, not ten. However, the quantiles all have different means in the (2005) online dataset, where it is stated that: "There is one difference compared to the [2002] EJ [Economic Journal] sample. Urban and rural expenditures in Egypt have been adjusted to reflect mean difference between urban and rural areas."

or groups of countries, where fifteen countries with large populations or economies (such as China, India, Italy, and the United States) are considered individually, and all other countries are clustered into eighteen country groups. For each country or country group they combine data on GDP per capita in PPP$ with income shares for eleven quantiles—the bottom nine deciles and the top two vigintiles (5 percent of population).[36] Thus their estimates of global inequality for each year are based on 363 (33 × 11) data points. Like Milanovic, they assume incomes to be equally distributed within each quantile. Unlike Milanovic, who takes incomes or expenditures directly from household surveys, they scale within-country distributions to per capita GDP, recognizing that "because of the obvious discrepancy between household purchasing power and GDP per capita, using GDP per capita in place of mean personal income may bias the estimation of the evolution of world inequality. Correcting for the share of non-household income in GDP or the share of non-consumption expenditures or taking into account the effects of changes in the terms of trade on the purchasing power of national agents proved impossible for the historical period. For comparability reasons, the GDP per capita convention was retained even after 1950, though a better approximation of international differences in mean living standards would have been possible" (Bourguignon and Morrisson, p. 730).

They find that inequality increases between 1820 and 1950, according to all measures, and that the subsequent trend varies by inequality measure (see our table 1). All indices except the standard deviation of log-income are higher in 1992 than in 1970. The income shares of the top quintile, decile, and vigintile increased uniformly from 1970 to 1992, the top decile increasing its share from 50.8 percent in 1970 to 53.4 percent in 1992. The share of the bottom 20 percent was the same in 1992 (2.2 percent) as in 1970, with a slight trough in 1980 (2.0 percent).

They decompose the Theil T and the mean log deviation (MLD) into between- and within-"country group" components, but their use of 33 country groups rather than individual countries to decompose inequality makes their decompositions not comparable with those in other studies.

5.3. *Sala-i-Martín (2006)*

Sala-i-Martín (2006) estimates global income distributions using within-country quintile shares scaled to per capita GDP in PPP$.[37] He uses quintile share data from Deininger and Squire (1996 updated), extended with UNU–WIDER data, and takes per capita GDP in PPP$ from the Penn World Tables 6.0 (Heston, Robert Summers, and Aten 2002, and known as PWT). He presents estimates for global inequality for each year between 1970 and 2000 based on observed and estimated data for 138 countries, representing 93 percent of the world's population in 2000. For those countries with survey data for more than one year, representing 84 percent of the world's population, he uses "a simple linear time-trend forecast" (p. 358) to fill in quintile shares for missing years. For countries with data for only one year he assumes a linear trend based on an average for "neighboring countries" (p. 359), defined as those belonging to the same World Bank region, which have surveys for more than one year. For countries with no survey data, he imputes average quintile shares and estimated trends from neighboring countries.

From these quintile shares, Sala-i-Martín constructs quintile means by scaling to per

[36] They use per capita GDP data from Maddison (1995) and describe filling in gaps in GDP and population data by applying "growth rates observed for comparable neighbouring countries over the same period" (Bourguignon and Morrisson 2002, p. 729). Income distribution data are obtained from a variety of sources. For countries without data, the "distribution was arbitrarily assumed to be the same as in a similar country for which some evidence was available for the appropriate period" (Bourguignon and Morrisson 2002, p. 730).

[37] Sala-i-Martín (2006) follows two previously circulated working papers, Sala-i-Martín (2002a, 2002b).

capita GDP from PWT. Using these quintile means he nonparametrically estimates a density function for each country-year using kernel density estimation (KDE) with a Gaussian kernel, and with the same bandwidth in every case. Each density function is normalized by population size, and these within-country distributions are aggregated to construct a world income distribution.[38] For every measure he finds inequality to be higher in 1980 than in 1970, but lower in 2000 than in 1980 (see our tables 1 and 2). For the variance of log-income, inequality is higher in 2000 than in 1970, but for the other measures it is lower.

Sala-i-Martín's smoothing technique is problematic. The theory of KDE, as described by Bernard W. Silverman (1986), is based on the assumption that the observations from which the density function is to be estimated are independently and identically distributed draws from the underlying distribution. However, the quintile means used by Sala-i-Martín are "trimmed means" (Stephen M. Stigler 1973) based on ordered income data and are, therefore, neither independently nor identically distributed. The existing theory of KDE (e.g., Silverman 1986) does not apply to quintile means.

There is also a problem with his use of a constant bandwidth for different countries. Sala-i-Martín (2006, p. 361) writes: "We follow the convention in the literature and use the bandwidth $w = 0.9 * sd * n^{-1/5}$, where sd is the standard deviation of log-income and n is the number of observations. Obviously, each country has a different sd so, if we use this formula for w, we would have to assume a different w for each country and year. Instead, we prefer to use the same bandwidth for all countries and periods. One reason is that, with a constant bandwidth it is very easy to visualize whether the variance of the distribution has increased or decreased over time. Given a bandwidth, the density function will have the regular hump (normal) shape when the variance of the distribution is relatively small. As the variance increases, the kernel density function starts displaying peaks and valleys." That is, more unequal distributions (as determined by the quintile means) will display more "peaks and valleys" than less unequal distributions. Yet more unequal distributions are not more likely to be multimodal than less unequal distributions. The supposed advantage that more unequal distributions are estimated with more peaks and valleys—which Silverman (1986, p. 15) describes as "spurious fine structure" due to a too-small bandwidth—simply means that such distributions are more poorly estimated. Inequality and multimodality (multipolarity) are independent concepts: a distribution with a thick upper tail is unequal, but not necessarily multimodal (e.g., the Pareto distribution).

While Sala-i-Martín's use of a constant bandwidth across datasets is erroneous, the question remains whether this makes a substantial difference when the distributions of countries are aggregated into a global distribution. Camelia Minoiu (2007) shows that it does indeed make a substantial difference. She does not test the measurement of global inequality, but she compares estimates of global poverty at various international PPP$ poverty lines using Sala-i-Martín's method with those derived from the use of alternative bandwidths, allowing these bandwidths to vary across countries according to the spread of data for each country. While Sala-i-Martín found 7.3 percent of the world's population living under $1-a-day (at 1985 prices) in 1990, Minoiu finds between 5.3 and 9.5 percent, depending on the bandwidth used. Sala-i-Martín's figure in 2000 was 5.7 percent, while Minoiu finds between 4.2 and 7.5 percent. At the $2-a-day poverty line, Minoiu's estimates range from 23.4 to

[38] The same estimation procedure is used by Sala-i-Martín (2002b) to construct a world income distribution, but in Sala-i-Martín (2002a) he pools all the quintile data points (for all 125 countries) and performs KDE across these to construct a world income distribution. The two procedures would generate the same world distribution were it not for the fact that different bandwidths are used in the two cases.

26.8 percent in 1990, compared with Sala-i-Martín's 16.2 percent, and from 15.0 to 20.7 percent in 2000, compared with Sala-i-Martín's 10.6 percent. Minoiu's dataset differs from Sala-i-Martín's so the estimates are not directly comparable, but the exercise demonstrates that estimates of poverty vary widely and are not robust to the use of different parameters. With such a wide range of estimates possible at different poverty lines, estimates of global inequality would also be expected to vary widely with bandwidth.

The final problem with Sala-i-Martín's procedure concerns the fact that his estimations are based on only five data points. As we have seen, the theory of KDE does not apply to quintile means so the standard results cannot be invoked. But it is at least worth noting that the use of KDE is intended for a far greater number of observations. Silverman (1986) produces a heuristic example based on seven observations, but notes that "it should be stressed that it is not usually appropriate to construct a density estimate from such a small sample" (p. 15). In the standard use of KDE the variance of the estimator is inversely proportional to the number of observations, so that a small sample implies a high variance (Silverman 1986, p. 39). As Adonis Yatchew (1998) notes, nonparametric techniques "require large (in some cases, astronomically large) data sets, since relationships are 'discovered' by examining nearby observations" (p. 672).

Sala-i-Martín defends his use of per capita GDP as the mean for within-country distributions, instead of using mean incomes from surveys, by observing that surveys are available for only few years. He objects that "we would have to somehow forecast these survey means for the missing country/year cells" (p. 357). But missing values for quintile shares for most years did not deter him from "forecasting" within-country distributions. More importantly, as Bourguignon and Morrisson (2002, p. 730) point out, GDP per capita has obvious failings as a measure of household income. We discussed the question of scaling within-country distributions to national accounts categories in section 4.1 above. As we saw, while the national accounts category of household consumption expenditure is different from household income, GDP is also different from household income as it includes several components of nonhousehold income.[39]

Sala-i-Martín (2006, p. 390, table IV) reports the within- and between-country components for the MLD and the Theil T index. For both measures, between-country inequality comprises about 70 percent of global inequality in 1970, falling to just over 60 percent in 2000. This is due to an absolute rise in within-country inequality and an absolute decline in the between-country component. Sala-i-Martín's (2006, p. 388) definition of the "within-country" component is "the amount of inequality that would exist in the world if all countries had the same income per capita (that is, the same distribution mean) but the actual within-country differences across individuals" (see also his note to table IV, p. 391). While this is correct for the within-country component of the MLD measure, it is *not* correct for that of the Theil T index, both of which are presented in his table IV. We discuss this further in section 6.1 below.

In addition to presenting his own analysis of global income inequality, Sala-i-Martín (2006, p. 382) comments on the discussion of the subject in the 2001 *Human Development Report* (HDR) of the United Nations Development Programme (UNDP 2001). He writes that it "argues that global income inequality has risen based on the following logic:

Claim 1: 'Income inequalities within countries have increased.'

Claim 2: 'Income inequalities across countries have increased.'

[39] Sala-i-Martín (2006, p. 357, footnote 5) cites Deaton (2005) in describing some of the disadvantages of using household consumption from national accounts. However, Deaton's point is that these are disadvantages relative to the use of survey means, and not relative to GDP.

Conclusion: 'Global income inequalities have also increased.'"

Sala-i-Martín points out that if Claim 2 refers to what we described as concept one inequality, which counts each country as a unit, then the conclusion does not follow, stating that "by adding up two different concepts of inequality to somehow analyze the evolution of world income inequality, the UNDP falls into the fallacy of comparing apples to oranges" (p. 382). However, UNDP (2001) does not make this argument, and makes no reference to changes in global interpersonal inequality.[40]

5.4. *Bhalla (2002)*

Bhalla (2002) constructs annual estimates of global inequality for income and consumption separately for each year during 1950–2000. He finds that the global income Gini increases from the late 1950s to the early 1970s (late 1970s for consumption), and then decreases until 2000 (2002, p. 174, figure 11.1).

Bhalla scales within-country distributions to per capita GDP for his measurement of global income inequality,[41] and to HFCE from national accounts for consumption inequality,[42] both measured at PPP (the sources of which are unclear—see below). According to his figure 11.1 on p. 174, his sources for within-country inequality are Deininger and Squire (1996); World Income Inequality Database (WIID, available at www.wider.unu.edu/wiid); World Bank, *World Development Indicators*, CD-ROM; Asian Development Bank (2002).[43] There is some confusion regarding the number of surveys he uses to construct his global inequality estimates. Table A.1 on p. 209 records that there are 317 surveys (income and expenditure) for the period 1950–80, and 604 for the period 1980–2000, for a total for 921. But in the text he writes that "construction of the dataset required the use of data for more than 1,000 household surveys" (2002, p. 38). Whatever the precise number may be, he has to impute within-country distributions for the majority of his 7,599 country-years (149 countries times 51 years). Moreover, there is concern regarding the quality of the surveys that he uses. Ravallion (2002) observes that only "about half of Bhalla's 600 distributions over 1980–2000 would pass the quality standards applied to the [World] Bank's calculations" (p. 8). Milanovic (2005, p. 208) counts 286 independent distributions in Bhalla's Appendix C, Table C.1 (pp. 218–23).

Like Sala-i-Martín (2006), Bhalla (2002) uses the quintile share data to estimate continuous within-country distributions, employing what he calls the "simple accounting procedure (SAP)" (p. 6). Whereas Sala-i-Martín (2006) uses nonparametric density estimation, Bhalla uses regression to fit a three-parameter Lorenz curve to the quintile shares (comprising four independent observations, since they add up to 1), using a functional form due to Kakwani (1980). However,

[40] It states that: "World inequality is very high" (UNDP 2001, p. 19). Sala-i-Martín (2006, p. 382, footnote 26) also refers to HDR 2003, but that publication states that trends in global income inequality are "ambiguous" (UNDP 2003, p. 39), not that they are rising.

[41] This is nowhere stated explicitly, as far as we can tell, but we deduce it from his comments that "published national accounts figures, provided the best basis for estimating world inequality" (Bhalla 2002, p. 173) and that "household income has to be approximated by per capita GDP" (Bhalla 2002, pp. 103–4, footnote 1). However, on p. 217 he states that his ratio of survey to national accounts mean uses per capita GNP.

[42] In fact, consumption distributions are constructed by scaling within-country distributions to 0.867 times HFCE from NA (2002, p. 128), but since this scaling is uniform across the world it makes no difference to his estimates of inequality (the deflation is for the purpose of estimating absolute income poverty).

[43] Yet on pages 212–13 he refers only to the first three and to the World Development Indicators website. Still elsewhere, on p. 208, he mentions that the Deininger and Squire (1996) and WIID datasets "have been supplemented by data available from the Web (World Bank poverty monitor, worldbank.org/research/povmonitor; and Milanovic's data on Eastern European countries), as well as data gathered for 18 Asian countries (Asian Development Bank 2002)." From such documentation it is unclear exactly which sources have been used by Bhalla to provide his within-country distributions.

he does not stop here. He states that: "The basic equation results are then *filtered* by SAP to satisfy the theoretical boundary constraints (i.e., the sum of the estimated shares of each quintile is actually equal to the *observed* shares, and the share of each percentile is equal to or larger than the share of the previous percentile). The filtering is done through an iterative procedure, whereby at the end of the first round, the shares of each individual percentile in the first quintile get estimated and fixed, then the next quintile, then the next, and so on. (The only somewhat 'arbitrary' and somewhat 'flexible' percentiles are the first and the last, and this flexibility shows up in the errors; see below.)" (Bhalla 2002, pp. 133–34, emphasis in original). We are unable to decipher exactly what the procedure entails.[44] If his object is to force the estimated Lorenz curve through the four observed cumulative quintile shares, then there are many ways to achieve this while satisfying his constraint that the curve be convex.[45] Moreover, it is not clear what role the initial fitted Lorenz curve is playing in this procedure. Given the inadequate documentation it is impossible to replicate his results independently, violating the first criterion for empirical research.

Bhalla (2002) reports accuracy tests of his estimation method against unit-level data from India and against published data on "selected percentiles, and the Gini" for the United States. For India he claims that "the SAP method is seen to be shockingly accurate. The constructed Ginis are within 1 percent of the true value in almost 90 percent of the cases" (p. 214). For the United States, he claims "the constructed and original Ginis are within a whisker of each other for all the years" (p. 134). He concludes that "the tests above suggest that the SAP method is accurate both at the aggregate Gini level (very, very accurate) and at the individual percentile level (very accurate)" (p. 134). Without knowing how the SAP method works it is not possible for us to comment on these tests. For all the reader can tell, the method could have been constructed in order to fit the U.S. and Indian data, implying that its accuracy in these cases tells us nothing about its potential accuracy in other cases. Ravallion (2002) writes "the fact that one specific Lorenz curve model gives a good fit for one country does not mean it will fit well for others. Indeed, we find that very different models of the distribution (either Lorenz curves or density estimation) are needed in different countries, and even different dates for the same country" (pp. 14–15).

After estimating within-country distributions, Bhalla scales these to per capita GDP and per capita HFCE measured in PPP$. His use of PPP sources is problematic. On p. 207 he reports using "Penn World Tables, 1985-base PPP prices, referred to as PWT 5.6; *WDI* 1998, which has PPP data, 1987 base, at both constant and current prices; PPP data, 1975 base, from Summers and Heston (1988), Heston and Summers (1991), referred to as HS; and IMF, *International Financial Statistics* CD-ROM, 2002" in addition to WDI (edition not specified) and Maddison (2001). These sources are inconsistent for two reasons. First, the sources use different methods for calculating PPP rates. For instance, recent World Bank estimates in the WDI follow the EKS method (Sultan Ahmad 2003), while PWT uses the GK method, as discussed above in section 4.2. Secondly, PPPs estimated for different base years are also inconsistent. PPPs are estimated in the International Comparison Programme (ICP) in a given year t. To calculate GDP in PPP$ in year $t+n$, one has to scale GDP in year t up or down by the country's real growth rate (nominal growth minus a price deflator).[46] GDP in PPP$ in

[44] This procedure is described even more opaquely in appendix B, p. 212.

[45] That is, "the share of each percentile is equal to or larger than the share of the previous percentile."

[46] In the case of PWT, there is a further stage of reconciliation after this updating.

year $t+n$ calculated in this manner can be very different from that obtained by use of an ICP conducted in year $t+n$.[47,48] Without more information on which source is used for which countries in which years it is impossible to infer the bias caused by Bhalla's confounding of sources.

Despite these and other questions that arise in relation to calculations of PPP exchange rates, Bhalla seems to disregard any controversy concerning PPP estimates. He writes: "No one—not the official source of poverty figures, or any institution, or any outside researcher—is questioning the PPP estimates. This is not because everyone believes that these figures are accurate; it is only because no one has the capacity, or the resources, to come up with a 'better' estimate of the PPP exchange rate" (2002, p. 94). Yet Sanjay Reddy and Thomas Pogge (2005) question the PPP estimates, and Dowrick and Akmal's (2005) Afriat measure is an attempt at an alternative PPP exchange rate. Earlier versions of both papers are cited in Bhalla's references. Finally, Bhalla does not appear to distinguish between GK and EKS PPPs.

In assessing the accuracy of his Simple Accounting Procedure, Bhalla comments: "Is there a particular bias in the SAP method? There cannot be, because, as the name suggests, the procedure is one of simple counting, and simple accounting" (2002, p. 181–82). The simplicity of the method will certainly elude the reader, as will the sense in which his global inequality calculations involve mere "accounting." Moreover, the lack of transparency regarding the method and sources preclude the possibility of judging the extent of bias. Finally, it should be noted that even simple procedures can be biased.

5.5. *Dowrick and Akmal (2005)*

This study follows the approach of Bourguignon and Morrisson (2002) and Sala-i-Martín (2002a) in pooling within-country quantiles, in this case quintile shares from Deininger and Squire (1996),[49] scaled up to GDP per capita. When country GDP is measured using the standard Geary-Khamis (GK) PPP rates in PWT, they find that all the measures they estimate (Gini, Theil T, squared coefficient of variation, variance of log-income) decrease from 1980 to 1993 (see our table 1). What is novel in their paper is the use of an alternative PPP exchange rate based on Afriat (1984), which they argue gives a better measure of comparative purchasing power across countries. When country GDP is measured using Afriat PPP rates, inequality increases by all measures over the period 1980–93. They also estimate global inequality at market exchange rates, which they find to be both considerably higher than when measured at either PPP rate, and to increase faster than the increase at Afriat PPPs.

Dowrick and Akmal also run simulations based on generated lognormal distributions to estimate by how much the assumption of equal incomes within quintiles understates inequality within a country. They find that the variance of log-income for data grouped by quintiles is 90 percent of the actual value, while grouping by deciles yields more than 95 percent of the actual value (2005, p. 224, figure 8). They conclude that "the quintile income shares that we and other researchers have used are likely to come close to

[47] When a country has more than one ICP survey, some sort of reconciliation procedure must be used; see section 7 below.

[48] Bhalla laments (2002, p. 96) that the World Bank's consumption PPPs—that is, PPP rates based on the consumption component of GDP—are available only for 1993, in contrast to their GDP PPP rates that are available for many years. But, as just described, GDP PPP rates are constructed by measuring relative prices in one year only, and then scaling up and down across years using domestic price deflators. He could therefore have followed the same procedure using 1993 consumption PPPs to construct consumption PPPs for other years.

[49] When only Ginis and not quintile share data were available in Deininger and Squire (1996), they estimate the single-parameter functional form for the Lorenz curve suggested by Chotikapanich (1993) and thereby obtain quintile shares.

capturing the full contribution of intracountry inequality to world inequality" (2005, p. 224). Supposing that within-country inequality accounts for 35 percent of global inequality, as measured by varlog, then a 10 percent underestimation of within-country inequality in each country would imply a 3.5 percent underestimation of global inequality. However, Sala-i-Martín estimates global inequality using pooled quintile shares from his sample of countries in his (2002a) paper, and smoothed within-country distributions in (2002b), both papers using the same dataset (which is different from his 2006 dataset). His estimate of the global variance of log-income in 1998 is 7.7 percent higher when he uses smoothed within-country distributions. Moreover, Dowrick and Akmal's experiment can tell us little about the impact on other measures of inequality.

5.6. *Dikhanov and Ward (2002)*

This study takes distributions of incomes and expenditures from Milanovic's (2002) dataset for "45 of the largest countries ... where reasonably consistent distributions were available for more than one reference year" (Dikhanov and Ward 2002, p. 6). They estimate smoothed within-country distributions by interpolating third-degree polynomials between observed points of the cumulative distribution function. They then scale these within-country distributions to what they refer to as the national accounts category of "personal consumption expenditure" from World Bank databases, converted into EKS PPP$.[50] Finally, "a global picture was built up by taking the available income distributions from the eight largest countries in each 'continental' region (for South Asia only five countries are used as the number of countries in the region is small and the five countries chosen comprise more than 90 percent of the total population) and filling the remaining gaps (about 1/6 in terms of global income and population) according to observed regional patterns" (Dikhanov and Ward, p. 6). They do not explain how they estimate values for missing years. In particular, they do not explain how they extrapolate backwards to 1970 from Milanovic's "benchmark" 1988 data and forwards to 1999 from his "benchmark" 1993 data.

The direction and magnitude of change in inequality depend on the inequality measure used (see our table 1). The income share of the bottom decile remains constant at 0.5 percent while that of the top decile increases slowly and steadily from 48.5 percent in 1970 to 54.3 percent in 1999. Despite the ambiguity in their measures, the authors conclude that "during the last three decades, the global income distribution became less equal (both between country and within country)" (Dikhanov and Ward, p. 12).

5.7. *Schultz (1998)*

Schultz (1998) uses GDP PPP data from the Penn World Tables 5.5, covering 120 countries with 93 percent of the world population in 1960 and 92 percent in 1989. He takes quintile shares for countries from Deininger and Squire (1996), using data only on those countries for which there are at least two nationally representative samples since 1950. This yields 509 observations across 56 countries. He then runs regressions on these observations with log per capita GDP, per capita GDP squared, year, and dummies for type of survey distribution and region in order to estimate within-country inequality for country-years without data.

He uses the variance of log-income (varlog) as his measure of inequality in order to construct global interpersonal inequality as the sum of within-country inequality and between-country inequality. However, there appears to be a problem with his procedure. The variance of log-income is decomposable into "between-country" and "within-country" inequality only if "between-country" inequality is calculated by assigning

[50] Dikhanov and Ward (2002) is the only study that neither scales survey means to GDP per capita nor uses GK PPPs, both of which we have argued should be avoided.

to everyone within a country the country's *geometric* mean income, not its *arithmetic* mean income (Anand 1983, pp. 201, 330–31). The "between-country" component thus calculated can then be added to the population-weighted within-country varlogs to give the global varlog. However, the "between-country" component calculated by Schultz is based on the per capita GDP of countries, i.e., their arithmetic mean incomes. He is evidently aware of this issue and points out that "the national income variable should refer to the mean of the logarithms of income" rather than "arithmetic mean income that is logged in this analysis of intercountry income inequality" (endnote 8). However, the problem remains, and the figures that Schultz reports do not refer to the global variance of log-income.

5.8. *Chotikapanich, Valenzuela, and Rao (1997)*

This study estimates lognormal income distributions for countries from Gini coefficients reported in Deininger and Squire (1996), which are then scaled to per capita GDP in PPP$ from PWT 5.6. Their sample of Gini coefficients comprises 36 countries, for which they estimate lognormal distributions for the years 1980, 1985, and 1990. However, Africa is represented by only Tunisia and Mauritius, comprising 1.5 percent of that continent's population. The virtual omission of sub-Saharan Africa, with about 10 percent of the world's population, is a major problem for their estimates of changes in global inequality. Since sub-Saharan Africa includes many of the poorest countries in the world, and the per capita GDP of this region fell by 10 percent over 1980–90 (*World Development Indicators Online*), its omission will lead to a downward bias in changes in measured global inequality.

5.9. *Korzeniewicz and Moran (1997)*

This study estimates global inequality in 1965 and 1992 at market exchange rates only. Korzeniewicz and Moran estimate global inequality using quintile shares for 46 countries, mostly from World Bank (1994), scaled to per capita GNP at market exchange rates. They estimate the Gini and Theil *T*, finding both to have risen (see our table 2). The income share of the poorest 30 percent declined dramatically from 2.1 to 1.0 percent, while the top 20 percent enjoyed a rise in their share from 82.0 to 88.9 percent.

6. *Decomposing Global Inequality*

6.1. *Inequality Between and Within Countries*

Many studies that estimate global income inequality "decompose" overall world inequality into between- and within-country components. Thus, Milanovic (2002) states that between-country inequality is 88 percent of global interpersonal inequality as measured by the Gini coefficient.

The impression conveyed by such "decompositions" is that some 80 percent to 90 percent of global inequality (depending on the measure and year) arises from differences in mean income between countries. An obvious and perhaps common understanding of such decompositions is that if between-country differences in mean income were eliminated (i.e., if concept 2 inequality were zero), but within-country inequality in each country were kept constant, then global inequality would only be some 10 percent to 20 percent of its measured value. Unfortunately, this is *not* the correct interpretation of the decompositions presented in the studies. Moreover, the meaning and relevance of what is presented is not always clear.

Doing the counterfactual exercise of eliminating between-country inequality but keeping within-country inequality constant in each country will generate a world income distribution with substantially more inequality than the implied residual in the Gini decomposition. For example, it is shown in Anand (1983, pp. 319–26) that the overall Gini is always greater than or equal to both a population-weighted average of subgroup

TABLE 4
WITHIN-COUNTRY INEQUALITY AT PPP EXCHANGE RATES

	1970	1980	1988	1990	1993	1998	1999	2000
Theil *T*								
Dikhanov and Ward (2002)[a]	0.211	0.221		0.230			0.267	
Dowrick and Akmal (2005) (GK)[b]		0.244			0.234			
Dowrick and Akmal (2005) (Afriat)[b]		0.274			0.289			
Sala-i-Martín (2006)	0.255	0.262	0.260	0.261	0.269	0.279	0.281	0.284
Theil *L* (Mean Log Deviation)								
Milanovic (2002)				0.194		0.224		
Milanovic (2005)				0.203		0.228	0.232	
Sala-i-Martín (2006)	0.246	0.256	0.273	0.278	0.290	0.310	0.315	0.319

Notes: [a] Dikhanov and Ward (2002, p. 11) report a decomposition of "the Theil," but the two components sum to the total for their "Theil index 2," not their "Theil index."
[b] Calculated from table 6, p. 223, as global inequality less the "between-country index," which, as we discuss in the text, is not the same as the income-weighted within-country component of the Theil *T* index.

Ginis and an income-weighted average of subgroup Ginis. Hence, the Gini coefficient of the hypothetical world income distribution where each country's mean income is equalized but relative inequality (the Gini) in each country is kept constant, will be at least as large as the population-share weighted average of country Ginis. Dowrick and Akmal (2003, p. 18), which is an earlier version of Dowrick and Akmal (2005), find the population-weighted average Gini across 47 countries, covering "over two thirds of the world's population," to be 0.364 in 1993. This is about 55 percent of the level of most estimates of the global interpersonal Gini (about 0.65), not 10 percent to 20 percent. Thus, within-country inequality according to this interpretation will account for at least 55 percent of global interpersonal inequality.

Dikhanov and Ward (2002), Dowrick and Akmal (2005), and Sala-i-Martín (2006) also decompose the Theil *T* index and estimate that between-country inequality accounts for between 64 percent and 76 percent of overall global interpersonal inequality. Unlike the Gini coefficient, the Theil *T* index *is* additively decomposable into between- and within-country components. However, the weights on the within-country Theil *T* indices are *income* and not population shares of the countries. Eliminating between-country inequality by equalizing the mean incomes of countries will therefore also change the measured within-country component: the elimination will leave a *population*-weighted average of the Theil *T* indices of countries, not the original income-weighted average (Anand 1983, p. 199). Like the Gini coefficient, the Theil *T* index thus also has a problem in interpretation of its between-country component. Of the inequality indices presented in the studies, only the Theil *L* measure (mean logarithmic deviation), which is additively decomposable with population-share weights, has a consistent interpretation of its between- and within-group components (see Anand 1983, pp. 198–202).

Restricting ourselves to the two "decomposable" Theil measures, all estimates but one find that within-country inequality has risen since 1970 (see our table 4). The exception is the Theil *T* GK PPP estimate of Dowrick

and Akmal (2005) for the period 1980–93. However, when Dowrick and Akmal refer to "intra-country inequality" they appear to mean global inequality less population-weighted between-country inequality (this is what we report in table 4). As we have just seen, this residual is not within-country inequality according to the decomposition of the Theil T index. Moreover, using Dowrick and Akmal's preferred Afriat PPP estimate, even this residual shows an increase.

The relatively uniform finding that within-country inequality has risen is also consistent with Cornia and Kiiski's (2001) analysis of the World Income Inequality Database, which covers 80 percent of the world's population and 91 percent of world GDP. They find that inequality has risen in the recent past in countries representing 59 percent of their sample population, and fallen in countries representing only 5 percent of their sample.

Considering between-country inequality, Dowrick and Akmal (2005) find the between-country component of the Theil T index to have declined from 1980 to 1997 using GK PPPs, but to have risen slightly from 1980 to 1993 using Afriat PPPs. Dikhanov and Ward's between-country component for the Theil T is at the same level in 1980 as in 1999, at 0.64, but lower in 1970 (0.61) and higher in 1990 (0.66), producing an ambiguous picture. Sala-i-Martín (2006) finds the between-country component for Theil L (MLD) to decline fairly steadily from about 1980 to 2000; for Theil T it has no obvious trend from 1970 to 1990, but then declines through to 2000. The decreases in global inequality using decomposable measures appear to be primarily due to declines in between-country inequality according to Dowrick and Akmal (2005) (for GK PPPs only) and to Sala-i-Martín.

6.2. *China and India*

China and India, with respectively 21 percent and 17 percent of the world's population (UNPOP 2002), are likely to be significant determinants of global inequality. China's growth rate has been substantially higher than the world average since 1977 and, given its low initial income, it could be expected to act as an equalizing force. To a lesser extent the same may be true of India, which has grown less fast than China but still faster than the world average since 1980.

While there is little doubt that per capita GDP in China has grown very fast over the last thirty years, there appears to be a scholarly consensus that official estimates overstate it (e.g., Maddison 1998). The estimates in Maddison (2001) and in PWT (see Heston 2001) show lower rates of growth than the official figures; all studies that use national accounts data use the growth rates of PWT or Maddison.[51]

Both Schultz (1998) and Sala-i-Martín (2006) test the extent to which China influences their results.[52] Schultz finds that without China, between-country inequality as measured by varlog would have risen during 1960–89 by 27 percent,[53] while excluding India makes little difference to the trend. Sala-i-Martín (2006) finds that without China global inequality rises from a Gini coefficient of 0.620 to 0.648 over 1970–2000, in contrast to the decline he finds with China.

While the exercise of excluding China or India is instructive from the point of view of accounting for global inequality and its evolution, it should be clear that it has no implications for global welfare. One cannot draw any conclusions about global welfare by use of a less-than-global sample (that is clearly unrepresentative). An alternative counterfactual is to include China as part of the globe, and to consider the implications for global inequality of different growth rates for China. If China were to grow at the same

[51] While Dikhanov and Ward (2002) use World Bank GDP and growth estimates for most countries, they report using growth rates from Maddison (2001) in "some cases, for example China" (Ibid.: 14).

[52] The effect on the level of global inequality of excluding China from the world (like undersampling the rich from a national distribution, as discussed above in section 4.1) is theoretically ambiguous.

[53] Calculation based on figures given in Schultz (1998, p. 322) and his table 1, p. 316.

rate as the rest of the world, it would make no contribution to changes in between-country inequality. But if it continues to grow faster than the world average growth rate, as it has done in the last three decades, it will (continue to) be an equalizing force for between-country inequality as long as its mean income is below the world mean. However, if China continues to grow faster than the world average once it reaches world mean income it will start to become a disequalizing force. Hence China's equalizing effect on between-country inequality may be temporary.

7. *Estimation Errors*

Most of the studies contain little discussion of potential sources of error in their estimates. Milanovic (2002), Bourguignon and Morrisson (2002), and Dowrick and Akmal (2005) all estimate standard errors of one kind or another, but in our view none of the studies accounts for the entire range of possible sources of error.

Milanovic (2002, p. 72) estimates standard errors in the Gini using the "jacknife" technique described in Arne Sandström, Jan H. Wretman, and Bertil Waldén (1988, p. 116), which is based on dropping each "observation" in turn and reestimating the Gini. The estimated standard errors for his common sample estimates are 0.031 in 1988 and 0.027 in 1993. Hence the confidence interval with two standard errors on either side of his estimate is (0.566, 0.690) for 1988, and (0.606, 0.714) for 1993. The measured change over the period would therefore appear to be insignificant. In Milanovic (2005), the corresponding confidence intervals for 1988 and 1998 are (0.586, 0.658) and (0.603, 0.679) respectively, and hence the change during 1988–98 is also insignificant. However, the "jacknife" technique is intended for estimating standard errors due to sampling variance—that is, the error arising from the sample not being representative of the population. But Milanovic's Gini coefficients are based on incomes estimated for all individuals in the world (in fact, 84 percent of them in his common sample). His world income distribution comprises deciles (or other quantiles) of countries' populations, whose incomes are estimated from surveys. While each survey is subject to sampling variance, his world income distribution is not based on a sample of the world population. In other words, he has a *constructed* population and not a *random sample* of the world population. Errors in his estimation procedure, therefore, are not based on sampling variance, and the interpretation of his estimated standard errors remains unclear. It should be noted that no other study even attempts to account for sampling variance explicitly.

Regarding estimation errors more generally, there are at least two levels at which data problems can arise in the studies that estimate global inequality. The first level is error in the data generated by the surveys and NA themselves. The second is measurement or estimation error in the PPP conversion rates used to construct a global distribution.

We have already discussed several sources of error in the NA and surveys. NA suffer from out-of-date sources for their benchmark data and unreliable imputations and extrapolations for current estimates. Surveys suffer from underreporting of the incomes and expenditures of the rich, and from sampling error, e.g., undersampling at both ends of the distribution. In addition to undersampling, the responses from those who are sampled will tend to be noisy. Even if noise does not affect the mean response, it is likely to influence the variance of responses. Hence, measured inequality could be overestimated. The extent of such noise or measurement error will vary across surveys, introducing another source of uncertainty in calculations of within-country inequality and hence in estimates of global inequality.[54]

[54] Over time, if this noise diminishes as survey methods improve, a decline in measured inequality might be expected simply on this score.

Bourguignon and Morrisson (2002, p. 730) simulate uncertainty in GDP figures and in within-country distributions, and estimate standard errors for global inequality on this basis. For GDP they assume that the data are normally distributed with "mean unity" (the mean of the multiplicative factor we presume) and a standard deviation of 10 percent during the nineteenth century, 5 percent during 1900–29, 2.5 percent during 1950–80, and 0 percent in 1992. For within-country distributions they calibrate stochastic errors in observed decile shares so that the resulting standard deviations of country Ginis average 2 Gini points in the nineteenth century and 1 Gini point in the twentieth century. Based on these assumptions, the resulting standard errors on the global Gini turn out to be small: in 1820 the standard error is 0.9 Gini points, in 1950 it is 0.2 Gini points (0.002 in the Gini scale of 0 to 1), and in 1992 it is 0.1 Gini points.

The small size of these estimated standard errors is a consequence of the assumptions about errors in the underlying data. Given the potential sources of error in the NA discussed above, these assumptions seem overoptimistic. A 2.5 percent standard deviation in the measurement of per capita GDP implies that its true value lies within 5 percent of its observed value in 95 percent of cases (including in underdeveloped regions of the world). Bourguignon and Morrisson's assumption of zero error in per capita GDP in 1992 seems even more optimistic. The errors assumed for the nineteenth century yield a 95 percent confidence interval of ±1.8 Gini points. If the same assumptions were made in the twentieth century and yielded the same confidence interval, then the estimated change in global inequality between 1950 and 1992 would be insignificant.

Apart from Milanovic (2002) and Bourguignon and Morrisson (2002), no other study seriously considers the implications of measurement error in within-country distributions. Schultz (1998) simply reports that his "estimates of intracountry household inequality are subject to a wide margin of error or uncertainty" (p. 326). Several other studies make reference to the unreliability of within-country estimates of inequality, but none investigates its implications.

In addition to errors due to sampling and to underreporting by the rich in household surveys, the assumption made in estimating within-country inequality of equal incomes within quantiles or income intervals is another source of error. This leads to a downward bias in within-country inequality in those studies that make this assumption—Milanovic, Bourguignon and Morrisson, Dowrick and Akmal, and Korzeniewicz and Moran. An analysis of the sensitivity of estimates to different assumed degrees of inequality within quantiles or income groups might be attempted. Specifically, it would be interesting to calculate by how much within-country inequality would increase if the distribution within income intervals were assumed to be maximally unequal. We could thereby construct an upper bound for within-country inequality in addition to the existing lower bound. The resulting "uncertainty" intervals would allow a better assessment of whether observed changes are significant.[55]

Only two studies, Bhalla (2002) and Sala-i-Martín (2006), construct explicit "smoothed" within-country distributions that do not assume equal incomes within population subgroups (or income intervals). However, we have seen that the methods employed by these studies are problematic. By contrast, there are well-developed procedures for estimating Lorenz curves from grouped data that allow reliable measurement of inequality—e.g., the World Bank's POVCAL software.[56] Of course,

[55] The assumption of zero inequality within income groups is as implausible as the assumption of maximal inequality.

[56] POVCAL was designed by Chen, Gaurav Datt, and Ravallion in 2001 and is described on the World Bank website at http://www.worldbank.org/research/povmonitor/software.htm. It involves the estimation of two distinct functional forms for the Lorenz curve—the Beta Lorenz curve and the General Quadratic Lorenz curve—discussed in Datt (1998).

it would be even better to obtain the primary (unit record) data from surveys and construct distributions from them directly rather than use secondary data on quintile shares. The World Bank appears to have a large collection of primary datasets which are utilized for estimating its global poverty figures, and indeed Milanovic (2002, 2005) has drawn on them to calculate global inequality.

In addition to error in the national accounts and in within-country distributions, measurement or estimation error will be present in PPP exchange rates. These exchange rates are estimated on the basis of data from price surveys undertaken in countries through the International Comparison Programme (ICP). PWT 5.6 had price data for only 85 of the 152 countries for which they present real income estimates, an increase over the 77 countries in PWT 5 for which they had price data (Summers and Heston 1991, p. 341). The latest version of PWT, i.e., version 6.2, has ICP price data for 115 countries. The countries with ICP price data are known as "benchmark countries." For nonbenchmark countries, PPP rates are estimated on the basis of regressions of indices of living costs in major cities carried out by a number of different organizations (Heston, Summers, and Aten 2004, pp. 3–4).[57]

PWT estimate PPP exchange rates for benchmark countries in nonbenchmark years using domestic price indices. However, as Aten and Heston (2004) report, "when countries have multiple benchmarks, the relative PPPs of two countries in two benchmarks usually differs from what would be predicted from relative price movements in the two countries" (p. 29). That is, suppose global ICP surveys are carried out in two benchmark years and used to calculate the PPP exchange rates between two countries in these years. As discussed above, the PPP exchange rate in the second benchmark year will not in general be equal to the exchange rate calculated by applying the two countries' domestic price indices to the PPP exchange rate in the first year.[58] Hence going forward from one benchmark year using domestic price indices will give a different result from going backwards from a subsequent benchmark year. Thus, for years in between benchmarks Aten and Heston "average the different PPP estimates and this is done by giving more recent estimates somewhat greater weight" (p. 29). Maddison and the World Bank do not follow this "reconciliation process" and there does not appear to be an accepted procedure for making PPP income comparisons over time. This introduces another layer of uncertainty in estimates of global inequality.

PPP exchange rates for benchmark countries will be subject to the standard sampling errors that arise in a (price) survey. For nonbenchmark countries, matters can be considerably worse. The authors of PWT 5 write that for nonbenchmark countries "the percentage accuracy, to be interpreted in 0.95 confidence interval terms, is guessed to range from 60 percent up or down for countries with GDPs per capita less than a tenth of the United States, to 19 percent up or down for countries between half and seven-tenths of the United States; and 15 percent for countries as close as seven-tenths of the United States" (Summers and Heston 1991, p. 341).

China and India pose particular problems for PPP comparisons. In the PPP data

[57] A preliminary report on the 2005 ICP, which will provide a new set of price data, appeared in December 2007 as this paper was going to press. In the 2005 ICP, country coverage has been much improved, with China and India included as benchmark countries. These price data are too recent to have yet been incorporated into PWT or WDI. In a privately-circulated note, Milanovic compares preliminary estimates of global inequality for 2002 based on the old (1993 ICP) and new (2005 ICP) PPPs. Using the 1993 ICP he finds the Gini for global inequality in 2002 to be 0.657, while using the 2005 ICP it is 0.699. This is a considerable increase, which is partly due to downward revisions in estimated income levels for China and India of approximately 40 percent. The effect that these revisions will have on measured changes over time in global inequality is unclear.

[58] This point has been emphasized by Reddy and Pogge (2005 and earlier drafts), and was discussed in a different context by Anand and Ravi Kanbur (1991).

used in the studies reviewed here, China is not a benchmark country,[59] while India was benchmarked only in 1985. Chinese PPP conversion factors in PWT are estimated on the basis of a regression equation. Dowrick and Akmal (2003, p. 21) examine the implications of this estimation for global inequality. They calculate a confidence interval for Chinese GDP in GK PPP$ (as in PWT), taking two standard errors on either side of the point estimate. Inserting the bounds of these estimated confidence intervals into their calculations of global inequality, they find that the resulting confidence intervals for global inequality in 1980 and 1993 overlap, and suggest that their estimated changes in global inequality are not robust to the estimation error in Chinese GDP.[60] Milanovic (2002, 2005) uses Chinese price data directly, but these data are based on surveys of only two cities in China and hence will be subject to large sampling error.

A detailed study by Deaton, Jed Friedman, and Vivi Alatas (2004) finds that both PWT and the World Bank substantially underestimate India's per capita GDP relative to that of Indonesia in 1999, in comparison to the authors' own bilateral PPP estimates. Their findings imply that India is at least 40 percent richer relative to Indonesia than PWT data would suggest, and at least 60 percent richer than the World Bank data would suggest.

Finally, there is the problem of assuming a single PPP price level for each country. If prices faced by households within a country are correlated with their incomes then assuming a single price level will bias estimates of within-country inequality upwards. Prices and incomes may be positively correlated if, for instance, both rural prices and incomes are lower than their urban counterparts. Not adjusting for price differences within the country will in this case lead to an upward bias in measured inequality. Conversely, income and prices may be negatively correlated if the poor face higher unit prices owing to their inability to buy in bulk (Aten and Heston 2004, pp. 7–9), which will lead to a downward bias in inequality. If relative prices within countries change over time then this may also lead to bias in estimated changes in inequality.

Evidently there are numerous sources of uncertainty—from errors in underlying data to biases arising from the assumptions and methods used to construct estimates of global inequality. Sensitivity analysis should be undertaken to assess the possible impact of these errors and assumptions, even if there is insufficient information to estimate statistical confidence intervals. The standard errors estimated in the literature do not address these concerns. If uncertainty intervals were constructed taking into account all these sources of possible error in estimates, they would in our view tend to render insignificant the measured changes in global inequality over the last thirty years of the twentieth century.

8. *Conclusion*

We have seen that studies of global inequality in general provide inadequate discussion of their methodology. Some studies are regrettably opaque and require considerable detective work to unearth the methods they employ. Even those that provide clear descriptions of their methods sometimes contain insufficient justification of the chosen method or discussion of its implications.

On the basis of our examination of the literature, we contend that it is not possible to reach a definitive conclusion regarding the direction of change in global inequality over the last three decades of the twentieth century. The different studies arrive at widely varying estimates of both the level of, and changes in, global interpersonal inequality.

[59] Price data are available for only two cities (Heston 2001; Milanovic 2002), which are not nationally representative.

[60] We would note, however, that the estimation errors may be positively correlated over time, in which case changes will be subject to less uncertainty than levels.

Estimates of the Gini in PPP$ over 1990 to 2000 range from 0.63 (Bhalla 2002) to 0.711 (Dowrick and Akmal 2005) while estimates of the Theil T range from 0.783 (Sala-i-Martín 2006) to 1.01 (Dowrick and Akmal 2005, using Afriat PPPs). Estimated changes over time range from a decline in the Gini of 0.04 over 1970–2000 (Bhalla 2002) to a rise in the Gini of 0.015 over 1970–99 (Dikhanov and Ward 2002). Bourguignon and Morrisson (2002) and Dikhanov and Ward (2002) find a rise in the Theil T from 1970 to the 1990s, while Sala-i-Martín (2006) finds a decline. Dowrick and Akmal (2005), studying the period 1980–93, find a rise using Afriat PPPs and a decline using GK PPPs. Even confining ourselves to estimates of the Gini using GK PPPs we find contradictory trends—from Sala-i-Martín's (2006) decline of 0.026 over 1970–2000 to Bourguignon and Morrisson's rise of 0.07 over 1970–92 (and no change over 1980–92). Finally, different inequality measures estimated for the same (constructed) global income distribution can imply contradictory trends, such as Dikhanov and Ward's estimate of a rise in the Theil T of 0.04 and a fall in the Theil L of 0.09 over 1980–99. The one point of agreement among all studies is that the level of global inequality is very high, with all estimates of the Gini since 1960 lying at or above 0.62. Most estimates also find that within-country inequality has risen since 1970 or 1980.

The diversity of the findings across studies is the result of varying data sources and methodologies. Studies use national accounts data from the Penn World Tables, Maddison (1995, 2001), or the World Bank—each of which employs a different PPP estimation methodology. Some assign a country its per capita GDP, some its per capita household consumption expenditure from NA, while others eschew national accounts altogether and use mean income or expenditure from household surveys (converted into PPP$ by one method or another). Within-country distributions may be constructed by assigning equal incomes within quintiles (or other quantiles), or may be "smoothed" using parametric or nonparametric methods.

We have provided arguments for preferring some choices over others in the many decisions that must be made in the construction of estimates of global inequality. We have argued that household surveys—the main source of estimates for within-country distribution—are more appropriate than national accounts for estimating the corresponding mean income. If within-country distributions must be scaled to a national accounts category then GDP is to be avoided, even if one is interested in household incomes rather than consumption expenditures. We also found that EKS PPPs are not subject to the biases suffered by market exchanges rates or by GK PPPs, and are more widely applicable than Afriat PPPs. More research is nonetheless warranted on PPP exchange rates for measuring global inequality, including the use of sector-specific and income-specific PPPs. In our view, two of the more serious problems are the scaling of incomes to per capita GDP, and the mixing of different and incommensurate PPPs.

Moreover, all studies suffer from a variety of sources of uncertainty that include inter alia: measurement error in national accounts, in household surveys, and in within-country price data used for PPP estimation (particularly serious for nonbenchmark countries including China); standard index number and multilateral comparison problems with PPP estimates; and noncomparability of household surveys. We do not know whether these errors will simply add noise or also lead to bias but, in either case, they reduce our confidence in measured changes in global interpersonal inequality.

Given these uncertainties, and the range of estimates for the direction and magnitude of change in global inequality, we conclude that there is insufficient evidence to reject the null hypothesis of no change in global interpersonal inequality over 1970–2000.

References

Ackland, Robert, Steve Dowrick, and Benoit Freyens. 2004. "Measuring Global Poverty: Why PPP Methods Matter." Paper presented at the 2004 Conference of the International Association for Research in Income and Wealth, Cork, Ireland, August. http://www.iariw.org/papers/2004/dowrick.pdf.

Afriat, Sydney N. 1981. "On the Constructability of Consistent Price Indices between Several Periods Simultaneously." In *Essays in the Theory and Measurement of Consumer Behaviour*, ed. A. Deaton. Cambridge: Cambridge University Press, 133–61.

Afriat, Sydney N. 1984. "The True Index." In *Demand, Equilibrium, and Trade: Essays in Honor of Ivor F. Pearce*, ed. A. Ingham and A. M. Ulph. New York: St. Martin's Press, 37–56.

Ahmad, Sultan. 2003. "Purchasing Power Parity (PPP) for International Comparison of Poverty: Sources and Methods." World Bank, Washington, D.C. http://siteresources.worldbank.org/ICPINT/Resources/poverty_PPP_Data_rev2.doc.

Anand, Sudhir. 1983. *Inequality and Poverty in Malaysia: Measurement and Decomposition*. Oxford and New York: Oxford University Press.

Anand, Sudhir, and Ravi Kanbur. 1991. "International Poverty Projections." PRE Working Paper WPS 617, World Bank, Washington, D.C., March; and Ld'A/QEH Working Paper No. 32, Queen Elizabeth House, Oxford, February.

Asian Development Bank. 2002. *Asian Drama Re-Visited: Policy Implications for the Poor*. Unpublished.

Aten, Bettina, and Alan Heston. 2004. "Use of Country Purchasing Power Parities for International Comparisons of Poverty Levels: Potential and Limitations." Paper presented at the 2004 Conference of the International Association for Research in Income and Wealth, Cork, Ireland, August. http://www.iariw.org/papers/2004/aten.pdf. Forthcoming in *Debates in the Measurement of Global Poverty*, ed. S. Anand, P. Segal, and J. Stiglitz. Oxford and New York: Oxford University Press.

Atkinson, Anthony B., and Andrea Brandolini. 2001. "Promise and Pitfalls in the Use of 'Secondary' Data-Sets: Income Inequality in OECD Countries as a Case Study." *Journal of Economic Literature*, 39(3): 771–99.

Atkinson, Anthony B., and Andrea Brandolini. 2004. "Global World Inequality: Absolute, Relative or Intermediate?" Paper Presented for the 28th General Conference of the International Association for Research in Income and Wealth, Cork, Ireland, August 22–28. http://www.iariw.org/papers/2004/brand.pdf.

Australian Treasury. 2001. "Global Poverty and Inequality in the 20th Century: Turning the Corner?" In *Economic Roundup Centenary Edition 2001*. http://www.treasury.gov.au/contentitem.asp?ContentID=110&NavID=.

Berry, Albert, François Bourguignon, and Christian Morrisson. 1983. "The Level of World Inequality: How Much Can One Say?" *Review of Income and Wealth*, 29(3): 217–41.

Bhalla, Surjit S. 2002. *Imagine There's No Country: Poverty, Inequality, and Growth in the Era of Globalization*. Washington, D.C.: Institute for International Economics.

Boltho, Andrea, and Gianni Toniolo. 1999. "The Assessment: The Twentieth Century—Achievements, Failures, Lessons." *Oxford Review of Economic Policy*, 15(4): 1–17.

Bourguignon, François, and Christian Morrisson. 2002. "Inequality among World Citizens: 1820–1992." *American Economic Review*, 92(4): 727–44.

Bureau of Economic Analysis. 2007. "A Guide to the National Income and Product Accounts of the United States." http://www.bea.gov/scb/pdf/misc/nipaguid.pdf.

Chen, Shaohua, Gaurav Datt, and Martin Ravallion. 1994. "Is Poverty Increasing in the Developing World?" *Review of Income and Wealth*, 40(4): 359–76.

Chen, Shaohua, and Martin Ravallion. 2001. "How Did the World's Poorest Fare in the 1990s?" *Review of Income and Wealth*, 47(3): 283–300.

Chen, Shaohua, and Martin Ravallion. 2004. "How Have the World's Poorest Fared since the Early 1980s?" *World Bank Research Observer*, 19(2): 141–69.

Chotikapanich, Duangkamon. 1993. "A Comparison of Alternative Functional Forms for the Lorenz Curve." *Economics Letters*, 41(2): 129–38.

Chotikapanich, Duangkamon, Rebecca Valenzuela, and D. S. Prasada Rao. 1997. "Global and Regional Inequality in the Distribution of Income: Estimation with Limited and Incomplete Data." *Empirical Economics*, 22(4): 533–46.

Cornia, Giovanni Andrea, with Sampsa Kiiski. 2001. "Trends in Income Distribution in the Post-World War II Period: Evidence and Interpretation." UNU–WIDER Discussion Paper, no. 2001/89.

Datt, Gaurav. 1998. "Computational Tools for Poverty Measurement and Analysis." International Food Policy Research Institute, Food Consumption and Nutrition Division Discussion Paper, no. 50.

Deaton, Angus. 2003. "How to Monitor Poverty for the Millennium Development Goals." *Journal of Human Development*, 4(3): 353–78.

Deaton, Angus. 2005. "Measuring Poverty in a Growing World (or Measuring Growth in a Poor World)." *Review of Economics and Statistics*, 87(1): 1–19.

Deaton, Angus, Jed Friedman, and Vivi Alatas. 2004. "Purchasing Power Parity Exchange Rates from Household Survey Data: India and Indonesia." Research Program in Development Studies, Princeton University. http://www.princeton.edu/~rpds/downloads/pppexchangerates.pdf.

Deaton, Angus, and Valerie Kozel, eds. 2005. *The Great Indian Poverty Debate*. New Delhi: MacMillan India.

Deininger, Klaus, and Lyn Squire. 1996. "A New Data Set Measuring Income Inequality." *World Bank Economic Review*, 10(3): 565–91.

Dikhanov, Yuri, and Michael Ward. 2002. "Evolution of the Global Distribution of Income, 1979–99." Paper Prepared for the 53rd Session of the International Statistical Institute held in Seoul, Republic

of Korea, August 22–29, 2001 and Updated for the 5th Conference on Globalization, Growth and (In)Equality held in Warwick, England, March 15–17, 2002.

Dowrick, Steve. 2002. "G–20 Comparisons of Incomes and Prices: What Can We Learn from the International Comparisons Program?" In *Globalisation, Living Standards and Inequality: Recent Progress and Continuing Challenges*, ed. D. Gruen, T. O'Brien, and J. Lawson. Sydney: Reserve Bank of Australia, 221–36.

Dowrick, Steve, and Muhammad Akmal. 2003. "Contradictory Trends in Global Income Inequality: A Tale of Two Biases." Paper presented at UNU/WIDER Conference on "Inequality, Poverty and Human Well-Being," Helsinki, Finland, May 2003.

Dowrick, Steve, and Muhammad Akmal. 2005. "Contradictory Trends in Global Income Inequality: A Tale of Two Biases." *Review of Income and Wealth*, 51(2): 201–29.

Dowrick, Steve, and John Quiggin. 1997. "True Measures of GDP and Convergence." *American Economic Review*, 87(1): 41–64.

Eltetö, O., and P. Köves. 1964. "On a Problem of Index Number Computation Relating to International Comparison." *Statisztikai Szemle*, 42: 507–18.

Firebaugh, Glenn. 1999. "Empirics of World Income Inequality." *American Journal of Sociology*, 104(6): 1597–1630.

Firebaugh, Glenn. 2003. *The New Geography of Global Income Inequality*. Cambridge and London: Harvard University Press.

Galbraith, James K., Lu Jiaqing, and William A. Darity, Jr. 1999. "Measuring the Evolution of Inequality in the Global Economy." University of Texas Inequality Project Working Paper, no. 7.

Grosh, Margaret E., and E. Wayne Nafziger. 1986. "The Computation of World Income Distribution." *Economic Development and Cultural Change*, 34(2): 347–59.

Gruen, David, Terry O'Brien, and Jeremy Lawson, eds. 2002. *Globalisation, Living Standards and Inequality: Recent Progress and Continuing Challenges*. Sydney: Reserve Bank of Australia.

Havinga, Ivo, Gisèle Kamanou, and Viet Vu. 2003. "A Note on the (Mis)Use of National Accounts for Estimation of Household Final Consumption Expenditures for Poverty Measures." Paper Presented at Global Poverty Workshop, Initiative for Policy Dialogue, Columbia University, New York, March–April.

Heston, Alan. 2001. "Treatment of China in PWT 6." http://pwt.econ.upenn.edu/Documentation/China.pdf.

Heston, Alan, Robert Summers, and Bettina Aten. 2001. "Penn World Table Version 6.0" Unpublished.

Heston, Alan, Robert Summers, and Bettina Aten. 2004. "Data Appendix for a Space-Time System of National Accounts: Penn World Table 6.1 (PWT 6.1)." http://pwt.econ.upenn.edu/Documentation/append61.pdf.

Kakwani, Nanak. 1980. *Income Inequality and Poverty: Methods of Estimation and Policy Applications*. Oxford and New York: Oxford University Press.

Korinek, Anton, Johan A. Mistiaen, and Martin Ravallion. 2006. "Survey Nonresponse and the Distribution of Income." *Journal of Economic Inequality*, 4(1): 33–55.

Korzeniewicz, Roberto P., and Timothy P. Moran. 1997. "World-Economic Trends in the Distribution of Income, 1965–1992." *American Journal of Sociology*, 102(4): 1000–1039.

Kulshreshtha, A. C., and Aloke Kar. 2005. "Estimates of Food Consumption Expenditure from Household Surveys and National Accounts." In *The Great Indian Poverty Debate*, ed. A. Deaton and V. Kozel. New Delhi: MacMillan India, 102–18.

Maddison, Angus. 1995. *Monitoring the World Economy: 1820-1992*. Development Centre Studies. Paris and Washington, D.C.: Organisation for Economic Co-operation and Development.

Maddison, Angus. 1998. *Chinese Economic Performance in the Long Run*. Development Centre Studies. Paris and Washington, D.C.: Organisation for Economic Co-operation and Development.

Maddison, Angus. 2001. *The World Economy: A Millennial Perspective*. Development Centre Seminars. Paris and Washington, D.C.: Organisation for Economic Co-operation and Development.

Melchior, Aren, Kjetil Telle, and Henrik Wiig. 2000. "Globalisation and Inequality: World Income Distribution and Living Standards, 1960–1998." Royal Norwegian Ministry of Foreign Affairs, Studies on Foreign Policy Issues Report, no. 6B.

Milanovic, Branko. 2002. "True World Income Distribution, 1988 and 1993: First Calculation Based on Household Surveys Alone." *Economic Journal*, 112(476): 51–92.

Milanovic, Branko. 2005. *Worlds Apart: Measuring International and Global Inequality*. Princeton and Oxford: Princeton University Press.

Minoiu, Camelia. 2007. "Poverty Analysis Based on Kernel Density Estimates from Grouped Data." Columbia University Institute for Social and Economic Research and Policy Working Paper, no. 07-07.

Nuxoll, Daniel A. 1994. "Differences in Relative Prices and International Differences in Growth Rates." *American Economic Review*, 84(5): 1423–36.

Pritchett, Lant. 1997. "Divergence, Big Time." *Journal of Economic Perspectives*, 11(3): 3–17.

Rao, D. S. Prasada. 2000. "Expenditure Share-Weighted Size-Neutral Geary–Khamis Method for International Comparisons: Specification and Properties." Unpublished. World Bank, Washington, D.C.

Rao, D. S. Prasada. 2003. "PPPs for the Measurement of Global and Regional Poverty: Issues and Options." World Bank, Washington, D.C. http://siteresources.worldbank.org/ICPINT/Resources/Rao_paper-PPPs_Measurement.doc.

Ravallion, Martin. 2000. "Should Poverty Measures Be Anchored to the National Accounts?" *Economic and Political Weekly*, 35(35–36): 3245–52.

Ravallion, Martin. 2002. "Have We Already Met the Millennium Development Goal for Poverty?" *Economic and Political Weekly*, 37(46): 4638–45.

Ravallion, Martin. 2004. "Competing Concepts of Inequality in the Globalization Debate." In

Brookings Trade Forum: 2004: Globalization, Poverty, and Inequality, ed. S. Collins and C. Graham. Washington, D.C.: Brookings Institution Press, 1–38.

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): 357–82.

Reddy, Sanjay, and Thomas Pogge. 2005. "How *Not* to Count the Poor." www.socialanalysis.org. Forthcoming in *Debates in the Measurement of Global Poverty*, ed. S. Anand, P. Segal, and J. Stiglitz. Oxford and New York: Oxford University Press.

Riskin, Carl. 2006. "The Fall in Chinese Poverty: Issues of Measurement, Incidence and Cause." In *Human Development in the Era of Globalization: Essays in Honor of Keith B. Griffin*, ed. J. K. Boyce et al. Cheltenham, U.K. and Northampton, Mass.: Elgar, 31–43.

Sala-i-Martín, Xavier. 2002a. "The Disturbing 'Rise' of Global Income Inequality." NBER Working Papers, no. 8904.

Sala-i-Martín, Xavier. 2002b. "The World Distribution of Income (Estimated from Individual Country Distributions)." NBER Working Papers, no. 8933.

Sala-i-Martín, Xavier. 2006. "The World Distribution of Income: Falling Poverty and . . . Convergence, Period." *Quarterly Journal of Economics*, 121(2): 351–97.

Sandström, Arne, Jan H. Wretman, and Bertil Waldén. 1988. "Variance Estimators of the Gini Coefficient–Probability Sampling." *Journal of Business and Economic Statistics*, 6(1): 113–19.

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(3): 307–44.

Silverman, Bernard W. 1986. *Density Estimation for Statistics and Data Analysis*. Monographs on Statistics and Applied Probability, vol. 26. Boca Raton, Fla. and London: Chapman and Hall/CRC.

Stigler, Stephen M. 1973. "The Asymptotic Distribution of the Trimmed Mean." *Annals of Statistics*, 1(3): 472–77.

Summers, Robert, and Alan Heston. 1984. "Improved International Comparisons of Real Product and Its Composition: 1950–1980." *Review of Income and Wealth*, 30(2): 207–62.

Summers, Robert, and Alan Heston. 1988. "A New Set of International Comparisons of Real Product and Price Levels: Estimates for 130 Countries, 1950–1985." *Review of Income and Wealth*, 34(1): 1–25.

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): 327–68.

Summers, Robert, Irving B. Kravis, and Alan Heston. 1981. "Inequality among Nations: 1950 and 1975." In *Disparities in Economic Development since the Industrial Revolution*, ed. P. Bairoch and M. Lévy-Leboyer. London: Macmillan, 18–28.

Summers, Robert, Irving B. Kravis, and Alan Heston. 1984. "Changes in the World Income Distribution." *Journal of Policy Modeling*, 6(2): 237–69.

Sundaram, K., and Suresh D. Tendulkar. 2003. "NAS–NSS Estimates of Private Consumption for Poverty Estimates: A Further Comparative Examination." *Economic and Political Weekly*, 38(4): 376–84.

System of National Accounts. 1993. Prepared under the Auspices of the Inter-Secretariat Working Group on National Accounts. Brussels/Luxembourg: Commission of the European Communities; Washington, D.C.: International Monetary Fund and World Bank; Paris: Organisation for Economic Co-operation and Development; New York: United Nations.

Szulc, B. 1964. "Indices for Multiregional Comparisons." *Przeglad Statystyczny*, 3: 239–54.

Theil, Henri. 1967. *Economics and Information Theory*. Amsterdam: North-Holland.

Theil, Henri. 1979. "World Income Inequality and Its Components." *Economics Letters*, 2(1): 99–102.

United Nations Development Programme. 1999. *Human Development Report, 1999*. New York and Oxford: Oxford University Press.

United Nations Development Programme. 2001. *Human Development Report 2001: Making New Technologies Work for Human Development*. Oxford and New York: Oxford University Press.

United Nations Development Programme. 2003. *Human Development Report 2003: Millennium Development Goals: A Compact among Nations to End Human Poverty*. Oxford and New York: Oxford University Press.

United Nations Population Division. 2002. *International Migration 2002*. New York and Geneva: United Nations Publications.

Wade, Robert H. 2001. "The Rising Inequality of World Income Distribution." *Finance and Development*, 38(4): 37–39.

World Bank. 1992. *World Development Report 1992*. Oxford and New York: Oxford University Press.

World Bank. 1994. *World Development Report, 1994: Infrastructure for Development*. Oxford and New York: Oxford University Press.

World Bank. 2001. *World Development Report 2000/2001: Attacking Poverty*. Oxford and New York: Oxford University Press.

World Bank. 2007. *World Development Indicators 2007*. Washington, D.C.: World Bank.

Yao, Shujie, and Liwei Zhu. 1998. "Understanding Income Inequality in China: A Multi-angle Perspective." *Economics of Planning*, 31(2–3): 133–50.

Yatchew, Adonis. 1998. "Nonparametric Regression Techniques in Economics." *Journal of Economic Literature*, 36(2): 669–721.