

CONTRADICTIONARY TRENDS IN GLOBAL INCOME INEQUALITY: A TALE OF TWO BIASES

BY STEVE DOWRICK*

School of Economics, Australian National University

AND

MUHAMMAD AKMAL

Australian Bureau of Agriculture and Resource Economics

Did global income inequality rise or fall over the last decades of the twentieth century? The answer depends on how cross-country income comparisons are made. Exchange rate comparisons suggest that inequality rose whilst the purchasing power comparisons of the Penn World Table suggest it fell. We show that both measures of real incomes lead to biased international income comparisons. Exchange rate comparisons ignore the relative price of non-tradables, whilst the fixed price method underlying the Penn World Table is subject to substitution bias. The contradictory trends are due to growing dissimilarity between national price structures increasing the degree of bias in each method. When we correct the income data to eliminate bias we find no compelling evidence of a significant change in world inequality.

1. INTRODUCTION

An apparent increase in inequality in international income distribution over the 1980s was highlighted by Korzeniewicz and Moran (1997) and by the United Nations Development Program in the Human Development Report (UNDP, 1999). Whilst the former is reluctant to draw strong conclusions about underlying causes, the UNDP report argues for international policies to mitigate rising inequality caused by economic globalization. Both studies rely on foreign exchange rates (FX) to convert incomes to a common currency.

For some purposes, such as assessing a nation's capacity to repay foreign debt or its bargaining power in international trade negotiations, the FX income comparison may well be appropriate. For the purpose of measuring inequality in living standards, however, we need to take account of the real purchasing power of national currencies which typically differs from the purchasing power implied by the exchange rate.

A number of studies¹ have shown world inequality to be falling when income comparisons are made using the purchasing power parities of the Penn World

Note: The authors thank the editor, Stephan Klasen, and two anonymous referees for their comments and suggestions. We are grateful for additional comments from participants at the WIDER 2003 Conference and from participants in seminars at Griffith University and the Australian National University.

*Correspondence to: Steve Dowrick, School of Economics, Copland Building (26), Australian National University, ACT 0200, Australia (Steve.Dowrick@anu.edu.au).

¹An exception is Milanovic (2002) who reports rising inequality between 1988 and 1993. Although Milanovic adjusts for purchasing power parities, his income definition is different from all the other studies, which are based on concepts of national or domestic income. He uses household income derived from survey data. Bourguignon and Morrisson (2002) find that the direction of change in inter-country inequality between 1980 and 1992 depends in part on the choice of inequality measure. Their Table 2 shows a small rise in the Theil index whilst both the mean logarithmic deviation and the standard deviation of log income decline. They use a different measure of purchasing power parity, relying on the estimates of Maddison (1995).

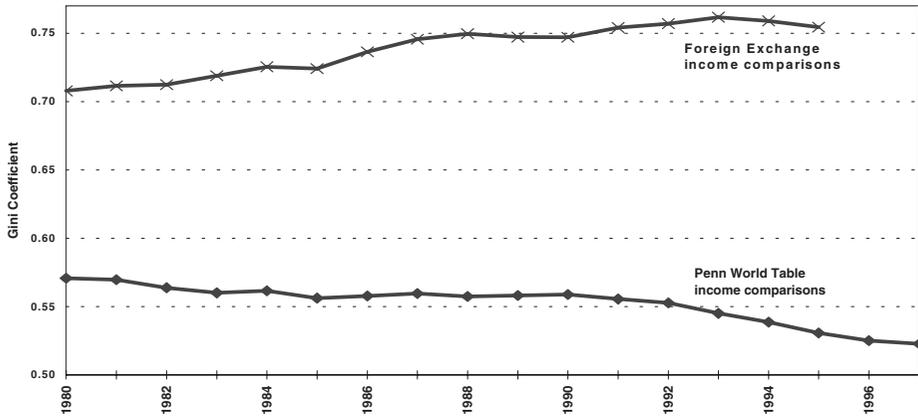


Figure 1. Inter-country Income Inequality 1980 to 1987

Source: World Bank (2001), Global Development Network database: Macro Time Series.

Notes:

1. “Foreign Exchange Income” is GDP at market prices (current US\$) divided by population. Dollar figures for GDP are converted from domestic currencies using single year official exchange rates.
2. “Penn World Table Income” is Real GDP Per Capita in constant dollars (international prices, base year 1985) derived from Penn World Table 5.6. Missing data are calculated by GDN from 1985 GDP per capita and GDP per capita growth rates using Global Development Finance and World Development Indicators.
3. Gini coefficients are calculated on *per capita* GDP for 115 countries using population weights.

Table (PWT5.6)²: Firebaugh (1999) reports that both the Theil and Gini indexes of inter-country inequality decline between 1965 and 1989, confirming the findings of Schulz (1998) who also reports a decrease in the variance of log income. These results are confirmed by Melchior *et al.* (2000), by Sala-i-Martin (2002b), and by our own calculations using PWT data, which are illustrated in Figure 1. We follow previous studies in using population-weighted measures of inequality, conceiving inequality as a function of income gaps between people or households. (If we focus on the national economy as the unit of observation, we find that PWT incomes have been diverging.)

It is well established that international currency markets tend to undervalue the domestic purchasing power of currencies of low productivity/low income countries. This is the common experience of international travelers who find that their money will go much further in India or Indonesia than it will in Western Europe. The phenomenon has been analyzed by Balassa (1964), Samuelson (1964) and Bhagwati (1984). Real wages are low in countries with low labor productivity, so non-traded labor-intensive services are cheap relative to capital-intensive traded goods. Market exchange rates are more likely to equate prices across countries in the traded sector than in the non-traded sector of the economy. Consequently, markets tend to undervalue the domestic purchasing power of the currencies of poor countries. An important implication of this “traded sector bias”

²See Summers and Heston (1991).

is that foreign exchange (FX) income comparisons, because they understate the real incomes of poorer economies, will overstate the degree of international inequality.

Alternatives to FX income comparisons rely on the estimation of purchasing power parities. A massive research effort, based on detailed price surveys in many countries under the auspices of the International Comparison Program (ICP), has resulted in the publication of the Penn World Table (PWT) which provides ready access to measures of *real* GDP per capita at constant international prices for over one hundred countries. These data are commonly referred to as PPP (purchasing power parity) measures of real income.

Many users of the PWT data are unaware, however, that attempts to measure purchasing power are problematic. The PWT are based on the Geary-Khamis (GK) method of construction of “average international prices” for a benchmark year. The GDP of each country in each year is valued at these fixed prices. However, constant price valuations introduce systematic bias by ignoring consumers’ ability to substitute towards goods and services that are locally cheap. The progenitors of the Penn World Table describe the problem of substitution bias thus:

The issue arises out of a familiar problem in price and quantity index number construction. . . . Valuation at other than own prices tends to inflate the aggregate value of the bundle of goods because no allowance is made for the substitutions in quantities toward the goods that are relatively cheap. . . . The practical importance of this issue . . . may loom large in comparisons between countries that have widely divergent price and quantity structures. (Kravis *et al.*, 1982, p. 7)

Use of the PWT estimates of international incomes, whilst avoiding the traded sector bias in the FX income data, introduces substitution bias in its place. We suspect that this issue may be very significant when it comes to assessing the level of and trends in world inequality, because we know that the price and quantity structures in the world’s poorest economies are very different from those typical of the richer industrialized economies, and we expect that price differentials vary over time.

Take, for example, the local currency prices for domestic services and for passenger cars, comparing an archetypal labor-intensive service with an archetypal traded good, in a sample of countries from the 1980 ICP survey. These prices are listed in the first two rows of Table 1. Local prices are for the real quantity that could be purchased for one Geary-Khamis international dollar.

The last row of Table 1 shows the price of domestic services as a ratio of the price of cars. We find that this price ratio varies by a factor of fifty between Germany and India. In Figure 2 we show a plot of this price ratio against *per capita* GDP for all of the 60 countries in the 1980 ICP survey, confirming that there is a very strong tendency for non-traded services to be much cheaper in low income countries than in rich countries.

Given this massive variation in relative prices, the problem of substitution bias is likely to be substantial. In Section 2 we develop a simple model that captures the biases in both the FX and GK measures of inequality. We find that whilst

TABLE 1
RELATIVE PRICES OF TRADED GOODS AND NON-TRADED SERVICES FOR FIVE COUNTRIES, 1980

	Canada	Germany	Brazil	Korea	Indonesia	India
Per capita GDP at Geary-Khamis international prices	\$11,573	\$10,390	\$3,758	\$2,393	\$1,101	\$605
Local currency prices						
Domestic services	1.44	3.23	7.68	116.2	70.4	0.45
Passenger cars	1.00	1.97	22.13	612.3	669.4	14.56
Price ratio						
Services/cars	1.44	1.63	0.35	0.19	0.11	0.03

Source: World Bank (1993), Tables 4 and 5.

Note: Local currency prices are defined as the number of units of a country's currency required to purchase the same amounts of goods and services as one U.S. dollar would purchase in the U.S.

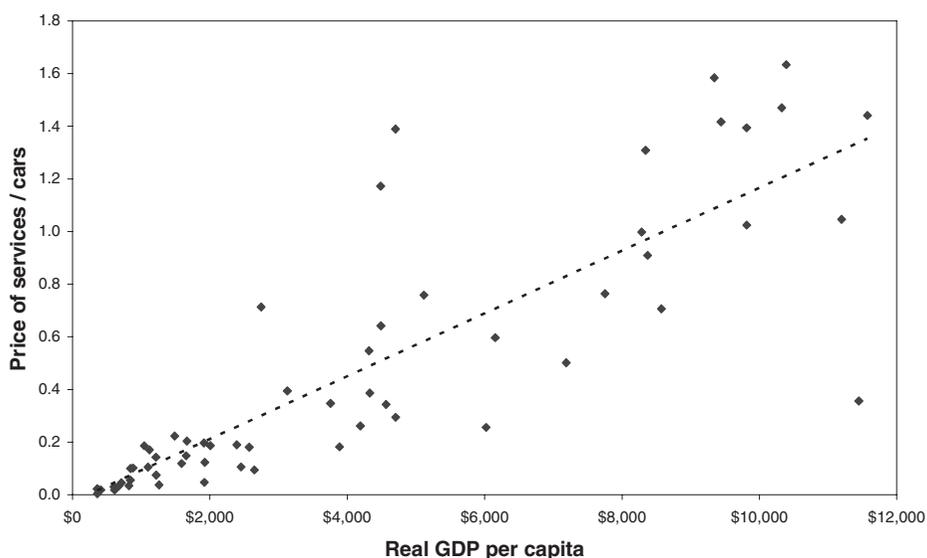


Figure 2. Price of Domestic Services Relative to Cars, 1980

Source: World Bank (1993).

Notes:

1. Real GDP *per capita* is measured in Geary-Khamis international dollars.
2. The price ratio is between the categories Domestic Services and Passenger Cars.

the Balassa-Samuelson effect causes FX comparisons to exaggerate inequality, the opposite bias is engendered by the GK method of measuring purchasing power parities. This suggests an explanation for the contradictory observations of trends in income inequality. If true inequality stays approximately constant, but measurement bias increases due to prices becoming less similar over time, then the FX method will show rising inequality whilst the GK method will indicate falling inequality.

We test this hypothesis in Sections 3 and 4. First, we investigate whether or not national price structures have diverged over the past few decades—a necessary condition for our hypothesis to hold. Second, we use the multilateral true index methodology of Dowrick and Quiggin (1997) to yield true PPP income comparisons that are free of both substitution bias and traded sector bias to test the conjecture that true inequality has been approximately constant.

2. A MODEL OF SECTORAL AND SUBSTITUTION BIAS IN INTERNATIONAL INCOME COMPARISONS

In order to better understand the sources of both FX and substitution bias, we construct a simple model of two trading economies. Each produces a non-traded labor-intensive service, S . Country 1 has comparative advantage in producing an intermediate good, A , which we might think of as an unprocessed agricultural or mineral product. Both countries manufacture a final tradable good, M , using labor and the intermediate good. The production technologies exhibit constant returns to scale and are identical across countries, except that country-specific private knowledge determines labor productivity in manufacturing. We define country 2 as the high-productivity country.

To keep the model simple we assume Cobb–Douglas production functions in manufacturing, we treat labor as the only factor of production, we disregard transport costs for trading the intermediate and manufactured goods and we assume competitive pricing behavior in product and labor markets, including free trade. We assume that all goods and services must be produced, traded and consumed within the one time period. The production side of the economy in country i can be summarized as follows:

$$(1) \quad S^i = L_s^i; \quad A^i = L_a^i; \quad M^i = (\lambda^i L_m^i)^\alpha \cdot (A_m^i)^{1-\alpha} \quad 0 < \alpha < 1$$

where $Z^i = (S^i, A^i, M^i)$ represent the domestic output of non-traded services, the traded intermediate product and the traded manufactured product respectively in country i ; L_z^i represents the amount of labor employed in each sector; A_m^i is the amount of intermediate input used in manufacturing; and λ^i is the productivity of labor in country i 's manufacturing sector.

Comparative advantage dictates that country 1 will export the intermediate good and import manufactures. We assume that the productivity differential and relative population size are such that it is feasible for country 1 to produce all of the intermediate good demanded in both countries.

Given constant returns to scale and competitive pricing, we can solve for the domestic price of the manufactured good in country i , P_m^i , in terms of the input prices for labor and the intermediate good, w^i and P_a^i (see Appendix for details):

$$(2) \quad P_m^i = \mu (w^i / \lambda^i)^\alpha (P_a^i)^{1-\alpha}; \quad \text{where } \mu = \alpha^{-\alpha} (1-\alpha)^{\alpha-1} \geq 1$$

We normalize prices and productivity by setting the wage and productivity level in country 1 to unity. We can then use λ (>1) without a superscript to repre-

sent manufacturing labor productivity in country 2. This allows us to derive the price vector for country 1 as:

$$(3) \quad \mathbf{P}^1 \equiv (P_s^1, P_a^1, P_m^1) = (1, 1, \mu)$$

The exchange rate is E units of currency 2 per unit of currency 1. The price of the imported intermediate good in country 2 is E . This determines the price of the manufactured good, using (2) as:

$$(4) \quad P_m^2 = \mu \left(\frac{w^2}{\lambda} \right)^\alpha E^{1-\alpha}$$

Free trade in the manufactured good equalizes prices across countries, requiring $P_m^2 = E \cdot P_m^1 = \mu E$. These conditions fully determine the wage in country 2. Setting the right hand side of (4) equal to μE yields:

$$(5) \quad w^2 = \lambda E$$

That is to say, productivity-adjusted factor-prices are equalized across the traded sectors.

By assumption, there are no differences across countries in the productivity of labor in the production of non-traded services. The price of services is simply the wage. It follows that services are relatively expensive in the high-productivity, high-wage country where the price vector, in units of country 1's currency, is:

$$(6) \quad \mathbf{P}^2 \equiv E(\lambda, 1, \mu)$$

We analyze demand and welfare by assuming common Cobb–Douglas preferences for the representative consumer-producer who is supplying a unit of labor inelastically:

$$(7) \quad U^i(s^i, m^i) = (s^i)^\beta (m^i)^{1-\beta} \quad 0 < \beta < 1$$

where s and m refer to *per capita* consumption of services and manufactured goods.

The budget share of services is β in each country. Given that *per capita* income in each country equals the wage, the *per capita* consumption bundles are:

$$(8) \quad \mathbf{q}^1 \equiv [s^1, m^1] = \left[\beta, \frac{1-\beta}{\mu} \right]; \quad \mathbf{q}^2 \equiv [s^2, m^2] = \left[\beta, \frac{\lambda(1-\beta)}{\mu} \right]$$

Per capita consumption of services is identical in the two countries, despite the fact that services are more expensive in country 2, because the income effect of higher manufacturing productivity offsets the price effect. This exact offsetting is an artifact of the Cobb–Douglas production and utility functions, but it is not crucial to our results. There is higher *per capita* consumption of manufactures in the higher productivity country. Thus we can refer to the higher productivity country as the high income or richer country.

Evaluating the common utility function (7) at \mathbf{q}^1 and \mathbf{q}^2 gives the welfare ranking $U^2 > U^1$. In general the utility ratio, U^2/U^1 , is greater than unity but otherwise indeterminate because the utility function is ordinal rather than cardinal. Cardinality is achieved by the money-metric Allen welfare index:

$$(9) \quad A_r^{2:1} \equiv \frac{e[U(\mathbf{q}^2), \mathbf{p}^r]}{e[U(\mathbf{q}^1), \mathbf{p}^r]}$$

which compares the minimum expenditures required to achieve utilities U^2 and U^1 at some reference price vector, \mathbf{p}^r . In our case, where the assumed preference function is homothetic, the Allen index is independent of the reference price vector. The true *per capita* income ratio between country 2 and country 1 is the unique value:

$$(10) \quad A^{2:1} \equiv \frac{U(\mathbf{q}^2)}{U(\mathbf{q}^1)} = \lambda^{1-\beta} > 1$$

These findings are summarized in the following proposition which is a restatement of the Balassa–Samuelson results:

Proposition 1: With free trade in intermediate and manufactured goods and competitive pricing, the country with higher productivity in manufacturing will exhibit the following features:

- (i) *per capita* utility is higher;
- (ii) non-traded labor-intensive services are more expensive relative to traded goods.

Foreign Exchange Rate Comparisons of Incomes

In this model, *per capita* National Income and Gross Domestic Product are identical and, measured in local currencies, are simply equal to the wage. So the GDP or income ratio that is obtained from exchange rate comparison is simply the ratio of the wage levels expressed in a common currency. Noting that the wage in country 1 is normalized to unity and substituting in equation (5) for wages in country 2, we derive the FX income ratio:

$$(11) \quad \text{FX}^{2:1} = \frac{w^2}{Ew^1} = \lambda$$

Proposition 2: Non-traded sector bias in FX comparisons

- (i) Market exchange rates overstate true international income differentials.
- (ii) The magnitude of the bias is an increasing function of: (a) the underlying productivity differential between the countries; and (b) the domestic expenditure share of the non-traded sector.

Proof: From (10) and (11), given $0 < \beta < 1$ and $\lambda > 1$,

$$(12) \quad \frac{\text{FX}^{2:1}}{A^{2:1}} = \lambda^\beta > 1$$

The first part of this proposition simply restates the more general result found in Samuelson (1974). The overstatement of true income differentials is due to “traded sector bias.” The market exchange rate equates purchasing power of currencies over the traded good but overstates the purchasing power of the currency over non-traded services in the rich country where non-traded labor-intensive ser-

vices are relatively expensive. It follows that the use of market exchange rates to compare incomes exaggerates the degree of international inequality. Our parameterization of the model leads to the unsurprising conclusion that the extent of this bias is increasing in both the productivity differential, λ (which causes the real wage differential) and the size of the non-traded sector, β (where the real wage differential affects the relative price).

Geary–Khamis Comparisons of Incomes

We turn now to the measurement of the international income ratio by the Geary–Khamis (GK) method which is used to construct the Penn World Table. This method values each country’s GDP bundle at “international prices.” The international price of manufactures, relative to services, is constructed as a quantity-weighted average of the relative prices of all the countries in the GK system. For the purposes of our model we have considered only two countries, but we can allow for other countries with a range of productivity levels in the GK system.

We represent the GK final consumption price vector as:

$$(13) \quad \mathbf{P}^{\text{GK}}(g) \equiv [P_S^{\text{GK}}, P_M^{\text{GK}}] = [g, \mu]$$

Referring back to equation (6), we see that the relative price of manufactures, μ/g , corresponds to the relative price that would be found in a country where the manufacturing productivity parameter is g . If the GK system is dominated by countries more productive than country 2, g will be greater than λ . If the rest of the world is less productive than country 1, g will be less than unity.

The Geary–Khamis measure of real GDP *per capita* for country i is the *per capita* consumption bundle evaluated at international prices: $(\mathbf{q}^i)' \mathbf{P}^{\text{GK}}$. Evaluating the consumption bundles given in (8) at prices $[g, \mu]$, the GK income ratio between countries 1 and 2 is:

$$(14) \quad \text{GK}^{2:1}(g) \equiv \frac{\text{GK}^2(g)}{\text{GK}^1(g)} = \frac{\beta g + (1 - \beta)\lambda}{\beta g + (1 - \beta)}$$

Whether this under- or over-states the true income ratio depends on the value of g . We summarize the relationship in our third proposition.

Proposition 3: Substitution bias in Geary–Khamis comparisons

- (i) A bilateral international comparison of *per capita* income which values expenditure at constant prices will understate the true income differential if the constant price vector corresponds to that of the high productivity country, or the prices of an even richer country.
- (ii) A constant price comparison will overstate the true income differential if the constant price vector corresponds to that of the low productivity country, or the prices of an even poorer country.
- (iii) The bias is greater, the less similar is the reference price vector with respect to the comparison country prices.
- (iv) Where (i) or (ii) holds, the magnitude of the bias is an increasing function of the underlying productivity differential between the two comparison countries.

Proof: See Appendix for details. The ratio of the constant price (GK) income ratio to the true income ratio is $R(g)$:

$$(15) \quad R(g) \equiv \frac{\text{GK}^{2:1}(g)}{A^{2:1}} = \frac{\beta g + (1 - \beta)\lambda}{[\beta g + (1 - \beta)]\lambda^{1-\beta}}$$

$R(g)$ measures the proportional bias in the GK index, with $R = 1$ representing no bias. Evaluating (15) gives $R(1) > 1$ and $R(\lambda) < 1$ (for $\lambda > 1$ and $0 < \beta < 1$, as assumed). Differentiating (15), $R_g < 0$. Hence (i) and (ii).

R is less than 1 for all $g > \lambda$. As g rises above λ (i.e. as the reference price vector becomes less similar to prices in countries 1 and 2) R falls. R is greater than 1 for all $g < \lambda$. As g falls below unity (i.e. as the reference price vector becomes less similar to prices in countries 1 and 2), R rises. Hence (iii).

R is decreasing/increasing in the productivity differential, λ , as $\lambda < g$ or $\lambda > g$. Hence (iv).

Proposition 3 clarifies the nature of substitution bias in fixed price comparisons. It is well known that the use of country 1's prices is likely to exaggerate country 2's welfare—since goods that are in high demand in 2, because of their relative cheapness, will be overvalued at 1's prices. In other words, the Laspeyres quantity index is usually larger than the Paasche index—and must be so if the underlying preferences are common and homothetic. It follows that valuing demand at country 1's prices will tend to overstate true inequality, if 1 is poorer than 2, and vice versa. This implication of substitution bias in the measurement of inequality is sometimes referred to as the *Gerschenkron Effect*, after Gerschenkron (1951).

Nuxoll (1994) has shown that the Gerschenkron effect will apply when the income ratio between country 1 and country 2 is measured at the prices of some third country, if relative prices and quantities are inversely correlated across all three countries. Proposition 3 formalizes Nuxoll's result in the context of an explicit model where prices, quantities and the true income ratio are endogenously determined by tastes and technology.

It follows from Proposition 3 that the direction and magnitude of bias in GK bilateral income ratios depends on whether the GK price vector corresponds most closely to the relative price structures of high income (high productivity) countries—in which case most bilateral ratios will be underestimated—or whether the GK price vector corresponds most closely to the relative price structures of low income (low productivity) countries—in which case most bilateral ratios will be overestimated. The former situation is most likely to apply given that the GK method weights each country's price vector by its share in total GDP, implying that more weight is given, *ceteris paribus*, to the price vectors of the richer countries.

The predicted magnitude of both fixed-price substitution bias and FX bias is illustrated in Figure 3 for the case where non-traded services comprise half of total expenditure, i.e. $\beta = 0.5$. The horizontal axis measures λ , the productivity differential between the manufacturing sectors of the two countries, which is illustrated for values between one and ten. The uppermost line shows the extent of positive bias in the foreign exchange comparison: $\text{FX}^{2:1}/A^{2:1}$. If the productivity differential is large, for example if $\lambda = 10$, FX comparisons overstates the true income ratio

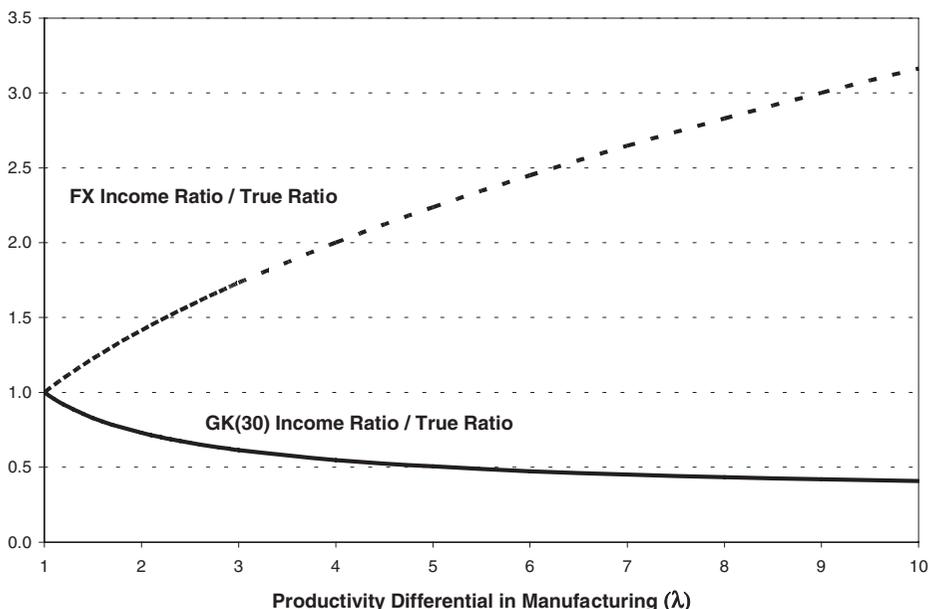


Figure 3. Predicted Exchange Rate Bias and Substitution Bias

Note: Predicted biases are calculated according to equations (12) and (15) with the parameters set as $\beta = 0.5$ and $g = 30$.

more than threefold. The lower line shows bias in GK income comparisons, $GK^{2:1}/A^{2:1}$, using constant prices that correspond to those of a high productivity country with a productivity parameter, g , of 30. In this case the GK method understates the true income ratio. This bias can be substantial; for example, it reaches a ratio of 0.4 when $\lambda = 10$.

The biases illustrated here refer to the measurement of the income ratio between a pair of countries. We expect to find similar bias in measures of multilateral inequality because most measures are constructed from bilateral ratios.

3. CAN WE EXPLAIN INCREASING FX INEQUALITY AND DECREASING PWT INEQUALITY?

We have demonstrated that there are potentially serious biases in the two most commonly used methods of comparing international incomes. Moreover, it is likely that the biases may work in opposite directions, with the Foreign Exchange (FX) income measure overstating the true level of inequality whilst the Geary–Khamis (GK) measure understates inequality. This occurs when the GK method is evaluating GDP at a set of “international” prices that correspond to the price structure of a high productivity economy.

Nuxoll (1994) analyzes bilateral income ratios and reports on page 1431 that the international prices underlying the 1980 ICP measures of real GDP correspond to those of “some moderately prosperous” country such as Hungary or

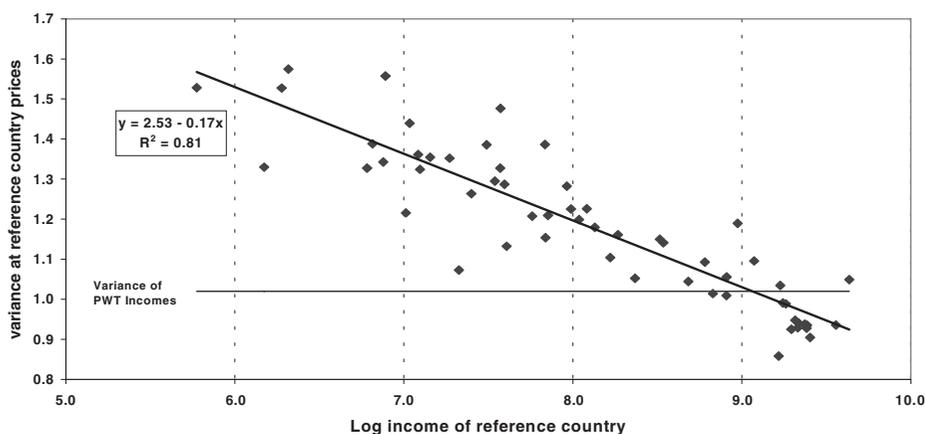


Figure 4. Variance of 1980 Log Real GDP *Per Capita* At Reference Country Prices

Source: World Bank (1993) Tables 4 and 5 for 1980 ICP data.

Notes:

1. “Income” is real GDP at Geary-Khamis international prices.
2. The “variance at reference country prices” is the variance of log GDP *per capita* measured for each country at the prices of the reference country.

Yugoslavia, with *per capita* GDP of around \$5,000. We apply a similar analysis to a multilateral measure, the variance of log income. We use the 1980 ICP data to construct measures of the variance of constant price log GDP *per capita* across the 60 countries, using first the prices of the U.S. and then the prices of each of the other 59 countries. The results are shown in Figure 4, with the variance on the vertical axis and the log of GK GDP *per capita* for the reference country on the horizontal axis. As predicted by our Proposition 3, the higher is the income of the reference country, the lower is the measured variance of log income across countries. The measured variances range from 0.85 to nearly 1.6. The empirical relationship is very strong with the linear regression line explaining more than 80 percent of the variation in measured inequality.

In Figure 4 we have also illustrated the variance of 1980 log incomes from PWT5.6 for the same countries, represented by the solid horizontal line at 1.02. This coincides with the regression line at a log income of 9.1, corresponding to income of $\exp(9.1) \approx \$9,000$. It appears that the GK constant international price vector underlying this version of the Penn World Table is most closely represented by the price structure of a relatively rich country such as Hong Kong, Japan or the U.K. with 1980 *per capita* income around \$9,000. This is not surprising, since the rich OECD economies, with average income levels above \$13,000, accounted for more than half of world GDP, exerting a dominant influence on the construction of the GK international price vector. More than two thirds of the countries have income levels lower than \$9,000, suggesting that most of the GK income ratios are likely to be biased downwards, understating the true level of inequality—as demonstrated by Dowrick and Quiggin (1997) and by Hill (2000).

We have shown that the degrees of bias in both FX and GK income measures are likely to increase as the true level of inequality increases. This suggests one

possible explanation for the conflicting messages coming from the analysis of FX and GK income inequality. If true inequality is rising over time, both the upward FX bias and the downward GK bias will increase. It follows that the rate of increase in inequality will be exaggerated by FX measures and understated by the GK measure.

On its own, this hypothesis is not enough to explain why FX inequality has been trending up and GK inequality trending down. In the simple model of Section 2, where there are only two consumption goods, both the FX and GK measures move in the same direction when true inequality changes. In reality, however, there is a multiplicity of goods and relative prices which are influenced to varying degrees by trade restrictions, by international differences in productivity and factor endowments, by monopoly pricing, by government regulation, etc. For instance, Falvey and Gemmell (1996) report that international differences in the price of services are explained by differences in factor endowments as well as differences in total factor productivity and Dowrick and Quiggin (1997) find that the degree of substitution bias in fixed-price income comparisons is increasing in the dissimilarity between countries' price vectors. It follows that the degree of bias in both FX and GK measures is likely to increase not only when underlying productivity differentials increase but also when national price structures diverge for other reasons. Thus we propose an explanation for the riddle of GK inequality falling whilst FX inequality rises:

Conjecture:

- (i) True inequality changed little between 1980 and 1993; but
- (ii) national price structures became less similar;

causing the FX measure of world inequality to rise and the GK measure of inequality to fall.

Firebaugh (1999) discounts this possibility, citing the finding of Dowrick and Quiggin (1997) that price structures had become more similar over the 1980s. Those findings were, however, only for the OECD countries where price convergence had been promoted by European economic integration. There is no presumption of price convergence across the rest of the world or between OECD countries and the rest.

We investigate trends in price similarity across the world using data from the Penn World Table 5.6 on the relative prices of private consumption, government consumption and investment for over one hundred countries. We use two measures of price similarity. First is the Kravis *et al.* (1982) definition of bilateral price similarity as the cosine of the angle between a pair of price vectors. For country i , we define its price similarity, PS_i^{GK} , in relation to the GK price vector P^{GK} as:

$$(16) \quad PS_i^{\text{GK}} = \frac{\sum_{j=1}^J w_j P_j^i P_j^{\text{GK}}}{\sqrt{\sum_{j=1}^J w_j (P_j^i)^2 \cdot \sum_{j=1}^J w_j (P_j^{\text{GK}})^2}}$$

with j indexing the categories of expenditure and w_j representing the share of total expenditure by all countries on that category.

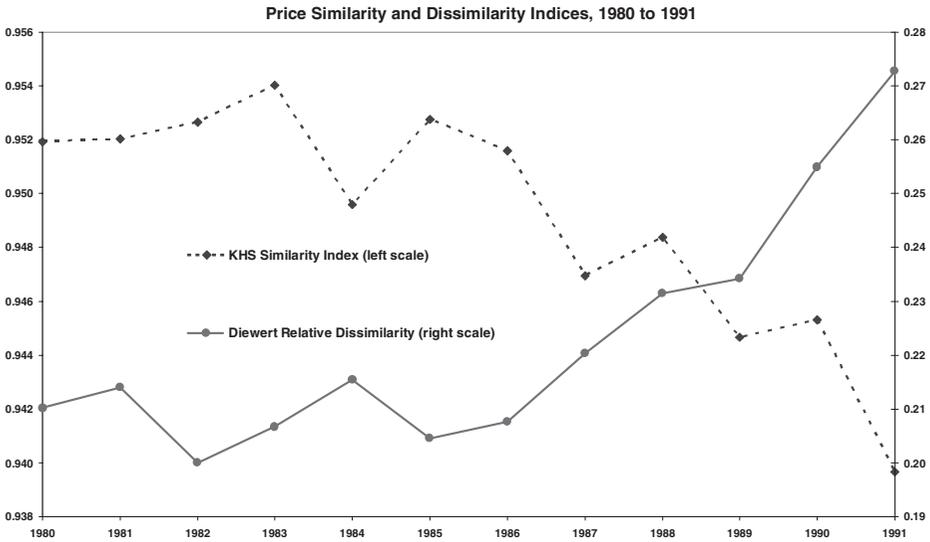


Figure 5. Price Similarity and Dissimilarity Indexes Across 131 Countries, 1980 to 1991

Source: Prices of private consumption, government consumption and investment are from Penn World Table 5.6.

Note: The index number formulae are equations (16) and (17). There are 131 country observations 1980 to 1989, 102 for 1990 and 100 for 1991.

Second, we use the Diewert (2002) bilateral index of relative price dissimilarity (the unweighted asymptotically linear index), $D(\mathbf{x}, \mathbf{g})$, which is defined as:

$$(17) \quad D(\mathbf{x}, \mathbf{g}) = \frac{1}{n} \sum_{i=1}^n \left(\frac{\mu g_i}{x_i} + \frac{x_i}{\mu g_i} - 2 \right)$$

$$\mu = \prod_{i=1}^n \left(\frac{x_i}{g_i} \right)^{\frac{1}{n}}$$

where \mathbf{x} represent the price vector of country x and \mathbf{g} represents the GK price vector, and $i = 1 \dots n$ indexes the commodities which make up the GDP bundle. Note that μ is defined as the geometric mean of \mathbf{x} relative to \mathbf{g} , and the dissimilarity index is the sum of the normalized price ratios and their reciprocals, adjusted to make the index equal zero when the prices in country x are an exact multiple of the GK prices.

Figure 5 displays the time trends between 1980 and 1991 for the cross-country means of the Diewert and KHS indices. The time trends are almost mirror images of each other because the former index measures dissimilarity rather than similarity. It is evident that between 1980 and 1991 price structures became markedly less similar/more dissimilar,³ confirming the second part of our conjecture.

³The sample of countries is held constant at 131 for 1980 to 1989, dropping to 112 and 100 in 1990 and 1991 respectively due to missing price data in PWT5.6. The results are very similar if we restrict our sample to the 98 countries for which data are available for all twelve years.

To investigate the other part of our conjecture we need measures of international incomes that are free of both substitution bias and traded sector bias. Dowrick and Quiggin (1997) develop a method that achieves this, based on the economic index number approach of Afriat (1984). This method involves searching for multilateral index numbers satisfying the condition that the bilateral ratios lie within the Paasche and Laspeyres bounds, making use of the algorithm described by Varian (1983). If such an index exists, the underlying price and quantity observations are consistent with optimizing choice by a representative consumer with homothetic preferences and the index numbers can be interpreted as an Allen welfare index. We apply this method to the 1980 and 1993 ICP benchmark studies.⁴

We find that for a majority of countries (52 of the 60 countries in the 1980 benchmark study, and 49 of the 53 countries in the 1993 study) the ICP observations on prices and quantities do satisfy the Afriat test for common homothetic preferences. For this set of observations the Allen money-metric utility comparisons, $E(u(q_i), p_r) / E(u(q_j), p_r)$, are independent of the choice of p_r . We then follow Dowrick and Quiggin (1997) in constructing the Ideal Afriat Index, which gives utility-consistent income ratios⁵ which are free of substitution bias, reflecting true purchasing power. We refer to this measure as Afriat income or, following Afriat (1984) and Dowrick and Quiggin (1997), as “true” income.

We note the finding of Ackland *et al.* (2004) that the EKS index is close to being a true Afriat index when applied to the 1993 ICP data. We confirm this finding for the 1980 ICP data where the variance of log income measured by the EKS index is 1.18, only slightly higher than the variance of log income measured by the Ideal Afriat Index. Indeed, a linear regression of the former index on the latter yields the result: $\ln \text{EKS} = 1.009 \ln \text{Afriat}$, $R^2 = 0.9989$. These results suggest that the subsequent data analysis would be altered very little if we were to substitute the EKS index for the Afriat index. We prefer to use the Afriat index because it is consistent with generalized homothetic preferences, as in our theoretical model, and because we are able to test whether the data are consistent with such preferences.

Estimating Afriat GDP for Non-Benchmark Countries

We are able to calculate true GDP per capita for all of the countries included in the 1980 and 1993 International Comparison (ICP) surveys, using the minimum of the Laspeyres valuations for those countries outside the homothetic sets as recommended by Dowrick and Quiggin (1997). However, in order to get genuinely representative estimates of world inequality, it is necessary to increase population coverage, requiring the extension of PPP estimates to non-benchmark countries.

⁴The 1993 ICP data have been made available by the World Bank (2000) on a regional basis. We understand that the Bank has not published the global estimates. We merged the regional data sets, using common countries to scale the price and quantity data, to create our own global estimates. We have also used ICP data published by World Bank (1993) and time-series data published on the Internet (World Bank, 2001).

⁵The Afriat method yields multiple sets of utility ratios with defined upper and lower bounds. For the purposes of this paper, we use the mid-points of the multilateral bounds, yielding the Ideal Afriat Index.

This is especially important because China is not covered in either of the ICP surveys and India is only in the 1980 benchmark.

In order to predict true incomes for non-benchmark countries, we use a procedure developed by Kravis *et al.* (1982)⁶ who use ICP benchmark data to estimate a regression model with GK income as the dependent variable and FX income as an explanatory variable.

Our model suggests that FX overstates true income for poorer countries. Substituting equation (11) into (10), and normalizing true income in country 1 to unity, implies that true income *per capita* in country i , A^i , is:

$$(18) \quad A^i = (FX^i)^{1-\beta} \Rightarrow \ln A^i = (1-\beta)\ln FX^i; \quad 0 < \beta < 1$$

We use this log-linear relationship as the basis of a regression model, augmenting it with a variable, OPEN, capturing the exposure of the country to foreign trade. Deviations between FX and A are driven by price differences between the traded and non-traded sectors of the economy—so where the traded sector constitutes a larger proportion of GDP, we expect the A/FX ratio to be closer to unity. We regress log A on log FX and OPEN for each of the two benchmark years separately, testing for functional form by including the square of log FX and the interaction of OPEN with log FX. Then, data for the two years is combined to obtain pooled estimates.

The regression results are reported in Table 2. Following tests for heteroscedasticity, the 1980 equation and the pooled model were estimated using weighted least squares whereas the 1993 equation was estimated by OLS.

The logarithm of FX income is highly significant in all equations. As expected, the coefficient is less than unity—varying between 0.6 and 0.8 according to the sample—implying that the true income distribution is less dispersed than the distribution of FX income.

The square of log FX adds significant explanatory power to the regression only in the 1980 sample. The openness variable is statistically significant in the 1993 regression, along with the interaction term, implying that there is less of a gap between FX income and true income in economies that are more exposed to world trade. The explanatory power, as defined by \bar{R}^2 , is close to 97 percent for the 1993 sample and the standard error of the regression is around 16 percent, which is very similar to the standard errors for the Kravis *et al.* (1982) regression reported in the final column of Table 2 and the regressions underlying the construction of PWT version 5.6 as reported in the December 1994 update to Appendix B of Summers and Heston (1991). The 1993 data are displayed in Figure 6 along with the predicted value of the regression. We see that a linear relationship between log A and log FX, as predicted by equation (18), fits the data reasonably well, with only minor additional explanatory power coming from the additional variables.

⁶Our approach to estimating real incomes in non-benchmark countries is similar to that reported in Appendix B for Summers and Heston (1991), except that they use various additional price surveys. An alternative approach used by Bergstrand (1991) uses the relative productivity levels in traded and non-traded sectors as an additional explanatory variable, but such data are not available for the non-benchmark countries.

TABLE 2

ESTIMATING THE RELATIONSHIP BETWEEN AFRIAT OR GEARY-KHAMIS INCOME AND FX INCOME

REGRESSION RESULTS

	1980	1993	Pooled Data	Geary-Khamis Comparison ³
Estimation by	Weighted Least Squares ⁴	Ordinary Least Squares	Weighted Least Squares ⁴	Ordinary Least Squares
Intercept	0.075 (1.36)	-0.137 (2.75)	-0.044 (0.70)	-0.086 (0.87)
D93			-0.121** (2.25)	
Log(FX)	0.795*** (27.1)	0.632*** (16.3)	0.742*** (16.3)	0.493*** (4.3)
[Log(FX)] ²	-0.045* (1.89)			-0.047* (1.84)
OPEN		0.200*** (2.71)	0.163* (1.77)	
OPEN × log(FX)		-0.104* (1.79)	-0.100 (1.53)	
Sample size	60	48	108	34
Standard error of regression	0.251	0.159	0.270	0.157
\bar{R}^2	0.936	0.968	0.915	0.967

Source: Authors' estimation columns 1-3. Kravis *et al.* (1982, p. 335), for column 4.

Notes:

1. The dependent variable is log Afriat GDP per capita in columns 1-3, and the Geary-Khamis measure in column 4.

2. t-statistics are given in brackets; ***indicates significance at the 1% level, **significance at the 5% level, *significance at the 10% level.

3. The regression summarized in column 4 is for 1975 Geary-Khamis measures of real GDP; it also contained statistically insignificant openness variables.

4. To take account of heteroscedasticity, the WLS weights are the predicted values of the absolute values of the residuals from the OLS regressions regressed on log FX.

DESCRIPTIVE STATISTICS

	Mean	St. Dev.	Min.	Max.
1980				
Log income (Afriat)	0.00	1.31	-2.65	2.07
Log income (FX)	0.00	1.09	-2.36	1.62
OPEN	0.62	0.32	0.12	1.81
1993				
Log income (Afriat)	0.00	1.52	-3.42	1.83
Log income (FX)	0.00	0.89	-2.20	1.20
OPEN	0.80	0.60	0.16	3.49

Source: World Bank (2000) for 1993 ICP data, World Bank (1993) for 1980 ICP data, and World Bank (2001) for FX data.

Note: Both FX and Afriat incomes have been normalized to a geometric mean of unity.

It is not clear why the OPEN variable is not significant in the 1980 sample, when we have a strong prior that the level of openness should be important in reducing the gap between FX and PPP incomes. One possible explanation is that we have misspecified the relationship between FX income and true income, perhaps through omission of variables or through incorrect functional form. This

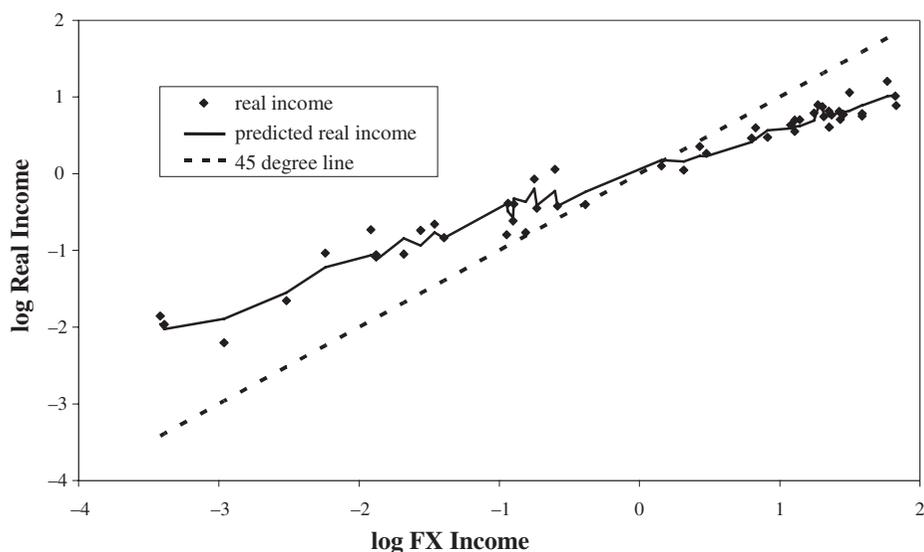


Figure 6. Real Afriat Income vs FX Income, 1993, Actual and Predicted Values

Source: World Bank (2001) for FX Income (GDP *per capita* at current exchange rates); World Bank (2000) for 1993 ICP data and authors' calculations of Real Income (Afriat index of true GDP *per capita*).

Notes:

1. "Predicted real income" is the predicted value of Afriat GDP *per capita* from the second regression reported in Table 2.
2. All of the inequality measures are population weighted as defined in equations (19) and (20).

poses a problem when it comes to predicting true incomes for the non-benchmark countries in each year. Should we use the particular coefficients for that year, or should we use the coefficients from the pooled regression? The former method may yield more accurate predictions for each particular year, but it creates further problems. The object of the exercise is to compare levels of true income inequality between 1980 and 1993. If we are using two different models to predict true incomes for non-benchmark countries, we cannot be sure whether any changes in estimated inequality are due to changes in the real income distribution or to the different methods used for predictions. Because we are primarily interested in the intertemporal comparison we focus on the results from the pooled regressions. We recognize that the possible misspecification of the exact relationship reduces the accuracy of the incomes predicted for non-benchmark countries, noting that similar problems apply to the PWT estimates of real income.

Different Measures of Inequality

We use four alternative measures of inequality—Gini (G), Theil (T), the squared coefficient of variation (CV^2) and the variance of logarithmic income (L). Firebaugh (1999) shows each measure can be represented by a distance function of the form:

$$(19) \quad I_m = \sum_{i=1}^N p_i f_m(y_i); \quad m = G, T, CV^2, L$$

where p_i is the share of the i -th country's population in the total population of all N countries and y_i is the ratio between the income per capita of the i -th country and the average income (per capita) across N countries. We use population weights to reflect our judgment that a proportional change in per capita income in a populous country such as China or India has more significance for global inequality than a similar change in a low population country such as Luxembourg or Iceland.

Firebaugh gives the specific functional forms that distinguish the four indexes:

$$(20) \quad G = \sum_{i=1}^N p_i y_i (q_i - Q_i); \quad T = \sum_{i=1}^N p_i y_i \log y_i; \quad CV^2 = \sum_{i=1}^N p_i (y_i - 1)^2;$$

$$L = \sum_{i=1}^N p_i \{\log y_i - E[\log y_i]\}^2,$$

where E is mean value operator, \log is the natural logarithm, q_i is the proportion of total population in the N countries that is poorer than country i and Q_i is the proportion of population richer than country i .

Intra-Country Inequality

Some studies of world income distribution have analyzed income inequality across countries while ignoring the intra-country component: Theil (1979, 1996), Theil and Seale (1994), and Firebaugh (1999). In other studies, however, the within-country dimension has been taken into consideration, giving a more accurate picture of inequality across all households in the world—as in Berry *et al.* (1983), Grosh and Nafziger (1986), Chotikapanich *et al.* (1997), Schulz (1998), Milanovic (2002), Sala-i-Martin (2002a, 2002b), and Bourguignon and Morrisson (2002).

Before considering within-country inequality we begin by analyzing the between-country component of global inequality for each of the three income measures: FX, Afriat and PWT. The inequality indexes cover 115 countries with 86 percent of world population in 1997. The PPP (Afriat) based indexes are calculated for 1980 and 1993 only.

Intra-country inequality measures have been taken from Deininger and Squire (1996) who have put together a data set containing quintile income shares and Gini coefficients classified by country, year, income type (gross or net), coverage (national or sub-national), form of recipient unit (person or household) and, importantly, by data quality. The data set includes 682 observations of the highest quality. Relying almost exclusively on the highest quality data, mostly on quintile distribution and occasionally on Gini, the four inequality indexes are computed for 1980 and 1993 for 47 countries.

For many countries Deininger and Squire do not report quintile distribution data of reliable quality for the two years. In order to increase the number of countries, the distribution data for the closest year, while constraining the departure to at most three years from the year of interest—1980 or 1993, are chosen as a proxy.

This increases the country coverage to 67, covering nearly 70 percent of the world population in both years.

In a few cases, only the Gini coefficient was available. In these cases we approximate the underlying quintile distribution using the single-parameter functional form of the Lorenz curve suggested by Chotikapanich (1993).⁷ The mean per capita income for different quintiles are obtained by multiplying the relevant quintile income share with the country's per capita income and then dividing by 0.2, the population share per quintile. We treat each country-quintile, with its average income and appropriate population weight, as a single observation in calculating global inequality.⁸

4. INEQUALITY RESULTS

The upper panel in Figure 7 repeats the illustration of FX and PWT Gini coefficients across 115 countries, but adds in our estimates of inequality of true Afriat incomes for 1980 and 1993. Our predictions of bias are confirmed. The FX measure overstates the true level inequality whilst the PWT measure understates it.

With respect to changes over time in the inequality of true incomes, there is a very slight rise in the inter-country Gini coefficient, from 0.615 in 1980 to 0.623 in 1993. This finding is not, however, robust to the inequality index employed. The remaining three panels of Figure 7 display the other indexes of inequality—Theil (7B), CV^2 (7C) and the variance of log income (7D). We see a slight reduction in the variance of log true income, from 1.551 to 1.522, between 1980 and 1993, whilst the other two measures of inequality register a slight increase.

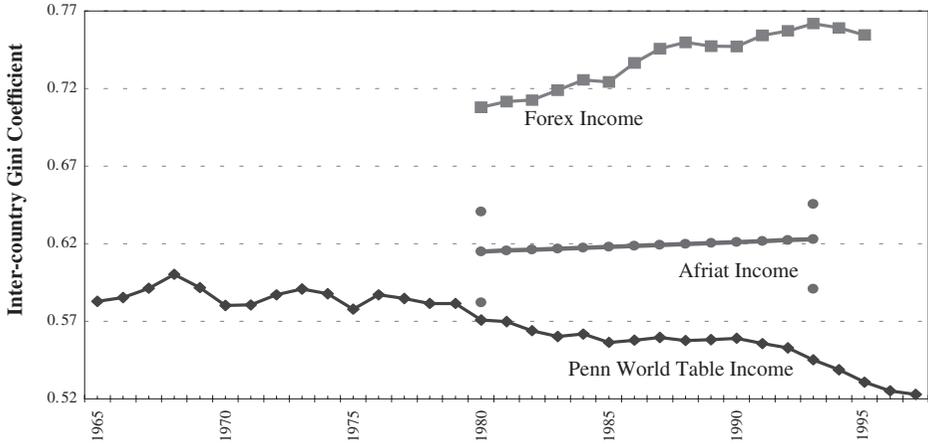
Estimation Error for Countries not Included in the ICP Benchmark Surveys

Because China was not included in either the 1980 or 1993 International Comparison Project price surveys, both the PWT and our own Afriat measures of real income rely on out of sample regression forecasts. Given that China accounts for over one-fifth of world population, it is crucial to analyze the robustness of measured inequality with respect to the prediction errors associated with Chinese real income. The standard errors of the regressions are at least 15 percent within the sample of benchmark countries—and we expect the error to be even greater when making predictions outside the sample. To check robustness, we calculate upper and lower bounds for real GDP *per capita* in China by adding or subtracting two standard errors from the baseline predicting regressions. The results are displayed

⁷Chotikapanich (1993) approximates the Lorenz curve by the following single-parameter specification: $LC = (e^{kp} - 1)/(e^k - 1)$. The corresponding Gini coefficient is given by: $G = [(k - 2)e^k + (k + 2)]/k(e^k - 1)$, where p is the population share and k is a parameter, which is required to be greater than zero. In the first step, the Gini equation is solved for k which then is used to obtain the estimates of quintile income distribution.

⁸These approximations, combined with our necessarily imperfect estimation of national incomes for non-benchmark countries and the more general problems of compiling data from secondary sources—see Atkinson and Brandolini (2001)—mean that there must be a substantial degree of imprecision in our inequality estimates. The principal interest of this paper, however, lies in comparing results from different methods of calculating purchasing power comparisons using the same secondary sources.

7A: Gini Coefficient



7B: Theil Index



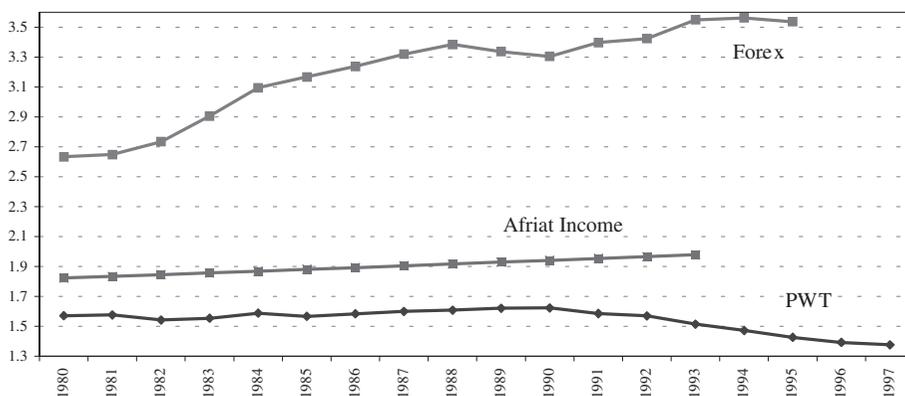
Figure 7. Inter-Country Inequality Across 115 Countries—Three Income Measures and Four Inequality Measures

Note: Afriat incomes are measured only in 1980 and 1993. The other series are annual. Dots around Afriat income represent lower and upper bounds when including Chinese income estimates.

in Table 3. In the case of the PWT income estimates, the upper and lower estimates suggest that *per capita* income in China in 1980 was between 5 and 9 percent of U.S. income. For the Afriat income estimates, the bounds are 3 and 8 percent.

The four inequality indexes are re-calculated for 1980 and 1993 using the upper and lower estimates of Chinese real GDP for both the Afriat and the PWT income measures. Results are reported in Table 4. The low and high values of the Gini differ significantly from their baseline counterparts. The “confidence intervals” for true income inequality are displayed as dots in the top panel of Figure 7.

7C: Squared Coefficient of Variation



7D: Variance of Log Income

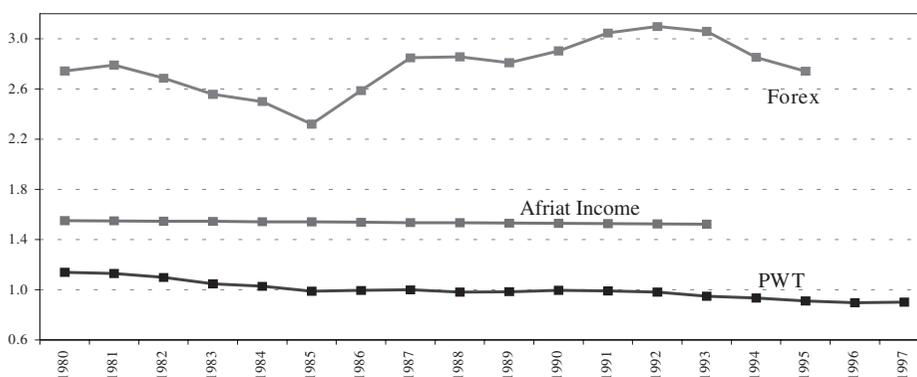


Figure 7. (continued)

TABLE 3
ESTIMATED GDP PER CAPITA OF CHINA (AS A PERCENTAGE OF US INCOME)

	1980			1993		
	FX	Afriat	PWT	FX	Afriat	PWT
Lower prediction		2.6%	4.7%		2.5%	6.7%
Baseline	1.7%	4.7%	6.4%	1.5%	4.5%	9.2%
Higher prediction		8.5%	8.7%		8.1%	12.6%

Notes:

1. The upper and lower predictions for Afriat income are calculated using forecast values from Regression 3 in Table 2 and adding or subtracting two standard errors of the regression.
2. The standard error of regression for the PWT income is taken from Summers and Heston (1991) who report seven alternative regression equations which are used for prediction depending on the availability of data on different countries. The median standard error of the regression, 0.15, is used to perform the sensitivity analysis.

TABLE 4
SENSITIVITY OF INTER-COUNTRY GINI COEFFICIENTS TO ESTIMATES OF CHINESE INCOME

	PWT Income			Afriat Income		
	Low Estimate	Baseline	High Estimate	Low Estimate	Baseline	High Estimate
1980	0.552	0.571	0.592	0.582	0.615	0.641
1993	0.526	0.545	0.566	0.591	0.623	0.646

Note: The low and high estimates of the Gini coefficients are calculated by replacing the baseline estimated income level for China with the high and low income estimates, respectively, from Table 3.

In another experiment we dropped China from our sample and computed inequality indexes for the rest of the 114 countries in the sample. We know that, as a country with relatively low income and relatively high growth, China is lowering global inequality over the time period. As expected, using either the FX or Afriat definitions of income, inequality increases more sharply when China is excluded. Using the PWT income estimates, the decline in world inequality disappears altogether when China is excluded.

We are not, of course, suggesting that China should be excluded from measures of inter-country inequality. But it is important to emphasize that all international comparisons of average real income for China are based on estimates, not on direct price measurement by the International Comparison Program. In the absence of such direct measurement, any estimates of recent trends in international inequality are subject to substantial uncertainty.

Global Inequality

Adding the dimension of within-nation inequality to our measures of international inequality allows us to analyze global inequality. We augment our measures of inter-country inequality, utilizing intra-country income distribution by quintiles, to derive estimates of global inequality for the 67 countries, including both China and India, for which sufficient data are available. Table 5 presents estimates of the global Gini coefficient for 1980 and 1993. We again find that global income inequality has been rising if we use the FX income definition, whilst the PWT measure of income records a fall. Our estimates of true (Afriat) income suggest a slight rise in global inequality.

In Table 5, the figures in parentheses give the ratio of the inter-country Gini to the global Gini calculated on quintile incomes for each country. This ratio varies between 85 and 89 percent for the PWT and Afriat measures of income and between 91 and 92 percent for the FX income-based Gini. Intra-country inequality contributes relatively little, no more than 15 percent, to the global Gini coefficient—in line with the findings of Berry *et al.* (1983), Milanovic (2002), Korzeniewicz and Moran (1997) and Li *et al.* (1998).⁹

⁹The relatively small contribution of within-country inequality is not surprising when we compare the quintile shares of income within and across countries in 1993. Within countries the ratio of income earned by the richest 20 percent is, on average, eight times the income earned by the poorest 20 percent of households. Across countries the inter-country quintile ratio of real *per capita* income is 25.

TABLE 5
GLOBAL INCOME INEQUALITY: GINI COEFFICIENTS BY COUNTRY
QUINTILES ACROSS 67 COUNTRIES

	PWT Income	Afriat Income	FX Income
1980	0.659 (86.7%)	0.698 (88.0%)	0.779 (90.9%)
1993	0.636 (85.6%)	0.711 (87.7%)	0.824 (92.4%)
Change 1980–93	–3.4%	1.7%	5.8%

Notes:

1. The average income of each country quintile is treated as a single observation. The Gini coefficient is calculated using country quintile population weights.

2. The numbers in parentheses are the inter-country Gini coefficients expressed as a percentage of the global coefficient.

TABLE 6
ALTERNATIVE MEASURES OF GLOBAL INCOME INEQUALITY BY COUNTRY QUINTILES ACROSS 67
COUNTRIES

	PWT	Afriat	FX
Theil Index			
1980	0.84 (70.9%)	0.96 (71.5%)	1.25 (77.3%)
1993	0.79 (70.4%)	1.01 (71.4%)	1.50 (79.0%)
Percent change 1980–93	–6.7%	+5.0%	+19.5%
Squared CV			
1980	2.86 (55.0%)	3.34 (54.5%)	4.33 (60.7%)
1993	2.73 (55.4%)	3.63 (54.5%)	5.73 (61.9%)
Percent change 1980–93	–4.4%	+8.7%	+32.1%
Variance of log			
1980	1.74 (65.3%)	2.21 (70.1%)	3.67 (74.7%)
1993	1.51 (62.8%)	2.40 (63.4%)	4.23 (72.2%)
Percent change 1980–93	–13.3%	+8.5%	+15.4%

Note: Values in brackets are the between-country index as a percentage of the global index.

Given that intra-country inequality is only a minor component of world inequality, it is not surprising to find that the global Gini coefficient displays a similar time trend to the inter-country Gini. The inclusion of within-country inequality tends to rescale the indexes upwards, without greatly altering rates of change over recent years.

The corresponding results for the other three inequality indexes—Theil (T), coefficient of variation squared (V^2) and variance of log income (L)—are presented in Table 6. We find that for each of these indexes the contribution of within-country inequality is proportionally greater than was the case with the Gini index. Nevertheless, the trends identified for the global Gini index, are essentially robust to the choice of inequality index. Global inequality falls between 1980 and 1993 for PWT incomes, but it increases for true Afriat income and FX income, according to all four indexes.

We are concerned, however, that the use of grouped income data may substantially understate the contribution of intra-country inequality to global inequality. As an experiment, we generated ten thousand log-normally distributed incomes and computed the variance of log income when the income is grouped

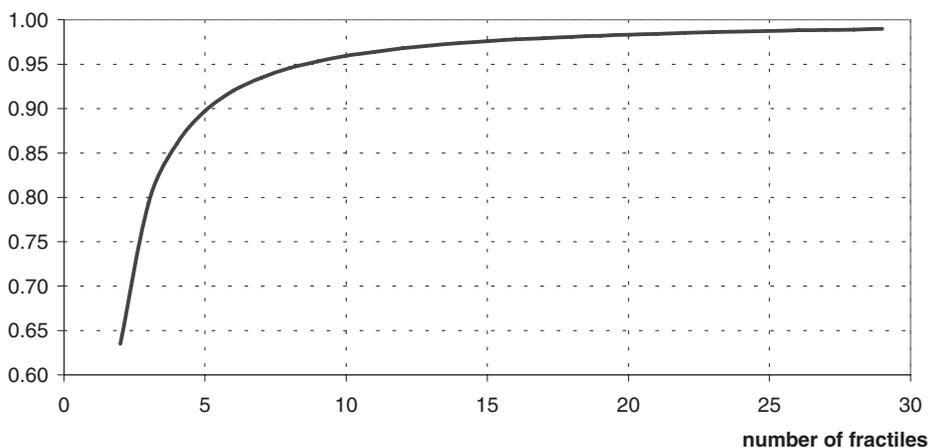


Figure 8. Variance of Fractile Income/Variance of Population

Note: Variances are calculated on 10,000 randomly generated log normal incomes which have been ordered and averaged over fractiles.

into various fractiles—as shown in Figure 8. We find that the variance of log income grouped into quintiles is 90 percent of the variance of the full sample. We conclude, therefore, that the quintile income shares that we and other researchers have used are likely to come close to capturing the full contribution of intra-country inequality to world inequality. This conclusion is reinforced by the finding of Sala-i-Martin (2002a) that estimating a Gaussian kernel density function has minimal impact on measures of global inequality.

Sala-i-Martin (2002b) criticizes our sample selection, based on an earlier version of this paper,¹⁰ on the grounds that “. . . the selection of countries that do not have Gini data is not random. In particular, these are countries that are poor and that have diverged. Excluding these countries from the analysis tends to bias the results towards finding reductions in world income inequality” (p. 9). This point simply reinforces our conclusion that, once we correct for the bias in the PWT measures of purchasing power parity, there is no convincing evidence of a fall in world income inequality over the period in question. Moreover, there is no such selection bias in our measures of the inter-country component which dominates quintile-based measures of global inequality.

5. CONCLUDING COMMENTS

A preliminary point that emerges from our theoretical and empirical analysis is that researchers who want to compare real income levels across countries need to be wary of the label “purchasing power parity.” There is no unique concept of purchasing power, and there are substantial differences in the methods under-

¹⁰Whilst criticizing our sample selection procedures, Sala-i-Martin ignores the principal point of our paper which is that the Penn World Table data, on which he is relying, bias the measurement of inequality.

lying the construction of widely used data sets which attach this label to their income measures. Our preference is to define purchasing power in terms of the capacity of a representative consumer to attain the same levels of utility when confronted with the particular price structure of each country. Using Afriat's non-parametric tests we are able to construct true income comparisons and to demonstrate that the frequently used Penn World Table data substantially understate the true level of income inequality due to substitution bias. We also note that the EKS index number approach, favored by the OECD in its calculations of purchasing power parities, does not suffer from such bias.

Regarding trends in global inequality, we replicate previous findings that the fixed-price method of calculating purchasing power parity incomes, which is used to construct the Penn World Table, leads to measures of inequality which tend to fall over the 1980s and 1990s, whilst market exchange rate comparisons of income suggest that inequality was rising. These observations raise a puzzle: with falling inequality in productivity and real income tending to reduce the sectoral bias of foreign exchange income comparisons, we should expect foreign exchange measures of income inequality to have been falling even faster than inequality measured at purchasing power parity. But the exact opposite has occurred, the gap between foreign exchange and PWT inequality has increased.

Our explanation for these contradictory trends in measures of global inequality rests on two hypotheses for which we have found empirical support. The first hypothesis is that both the foreign exchange method and the Penn World Table's fixed-price method of comparing incomes across countries are biased, the former method tending to overstate inequality and the latter method tending to understate it. These are the predictions of the trade model that we develop in Section 2. Our model exhibits the standard Balassa–Samuelson effect whereby cross-country differences in productivity in the traded sector lead to lower relative prices in the non-traded sectors of low productivity countries, leading to an exaggeration of true income differentials when national incomes are compared at the market rate of foreign exchange. The novel result of our model is that these intersectoral price differentials impart a downward bias to fixed-price measures of income differentials when the price vector is similar to that of high productivity countries. Our empirical analysis suggests that the fixed international price vector underlying the valuations of the Penn World Table does indeed correspond to the price structures of relatively rich economies.

The second hypothesis is that national price structures became increasingly dissimilar over recent decades. The biases in both PWT and foreign exchange measures of inequality are driven by differences in relative prices across countries—if relative prices were identical in all countries, there would be no bias in either method of valuing real incomes. Our model demonstrates that the magnitude of the bias is increasing in the size of the sectoral price differentials. As price structures become less similar, there is increasing downward bias in the PWT estimates of income inequality and increasing upward bias in foreign exchange estimates.

Over a period when world trade has been increasing, it may seem unlikely that national price structures would have become less similar—but this is exactly what we find when we examine trends in indexes of cross-country price similarity and dissimilarity. Such a result is not necessarily contrary to the Balassa–Samuelson

model where the bias in foreign exchange valuations of income arises out of the underlying productivity differentials and the relative sizes of the tradable and non-tradable sectors. An increase in actual trade will not affect the bias if the relative sizes of the domestic production sectors are unchanged, but an increase in the productivity differential will increase the bias. Moreover, a wide range of domestic supply and demand factors, as well as changes in government tax and subsidy policies, can be expected to influence domestic price structures differently in different countries. Further investigation of these issues is clearly warranted.

Here we have an explanation for the radical differences in measured inequality trends. True inequality was stable or increasing slightly—as suggested by our measures of true Afriat income—over the 1980s and 1990s whilst price structures became increasingly dissimilar.

Whilst this explanation is plausible, we cannot be sure that it is true. There are substantial errors involved in estimating real incomes for countries that have not been included in the International Comparison Program benchmark surveys, errors which apply equally to our estimates of Afriat incomes and to the Penn World Table income estimates. Moreover, until the publication of the 2004 ICP survey, which for the first time includes both India and China, all methods of calculating purchasing power parities have to resort to imprecise estimates of real income for more than one third of the world's population.

APPENDIX: MATHEMATICAL PROOFS

Derivation of Equation (2)

Output in sector m of country i is given by the production function

$$(A.21) \quad M^i = (\lambda L_m^i)^\alpha (A_m^i)^{1-\alpha}$$

Assuming competitive markets, the wage w^i and the price of the intermediate input, P^{ia} , are equated with the value of their marginal products:

$$(A.22) \quad P^{im} \alpha \lambda^\alpha (L^{im})^{\alpha-1} (A^{im})^{1-\alpha} = w^i \Rightarrow \left(\frac{L^{im}}{A^{im}} \right)^{1-\alpha} = \frac{\alpha \lambda^\alpha P^{im}}{w^i}$$

$$(A.23) \quad P^{im} (1-\alpha) (\lambda L^{im})^\alpha (A^{im})^{-\alpha} = P^{iA} \Rightarrow \left(\frac{L^{im}}{A^{im}} \right)^\alpha = \frac{P^{iA}}{(1-\alpha) \lambda^\alpha P^{im}}$$

Equating relative factor demands from (A22) and (A23) gives result (2):

$$(A.24) \quad P^{im} = \frac{1}{\alpha^\alpha (1-\alpha)^{1-\alpha}} \left(\frac{w^i}{\lambda} \right)^\alpha (P^{iA})^{1-\alpha}$$

Proof of Proposition 3

The GK income ratio, expressed as a proportion of the true income ratio, is $R(g)$:

$$(A.25) \quad R(g) = \frac{\beta g + (1-\beta)\lambda}{[\beta g + (1-\beta)]\lambda^{1-\beta}}$$

Differentiating with respect to g , representing the denominator of (A.25) as D , gives:

$$(A.26) \quad \begin{aligned} D^2 R'(g) &= \beta(1-\beta)(1-\lambda)\lambda^{\beta-1} \\ &\Rightarrow R'(g) < 0 \quad \forall \lambda > 1, 0 < \beta < 1 \end{aligned}$$

We need to show that $R(1) > 1$ and that $R(\lambda) < 1$.

$$(A.27) \quad R(1) = \frac{\beta + (1-\beta)\lambda}{\lambda^{1-\beta}} > 1 \quad \text{iff} \quad \beta + (1-\beta)\lambda - \lambda^{1-\beta} > 0$$

$$(A.28) \quad R(\lambda) = \frac{\lambda^\beta}{\beta\lambda + (1-\beta)} < 1 \quad \text{iff} \quad (1-\beta) + \beta\lambda - \lambda^\beta > 0$$

The conditions in (A.27) and (A.28) are of the same form:

$$(A.29) \quad \begin{aligned} f(\lambda) &\equiv d + (1-d)\lambda - \lambda^{1-d} > 0 \\ &\text{where } 0 < d < 1, \text{ since } d = \beta \text{ or } d = (1-\beta) \end{aligned}$$

We evaluate the function $f(\lambda)$ as follows:

$$(A.30) \quad \begin{aligned} f(1) &= d + 1 - d - 1 = 0 \\ f'(\lambda) &= (1-d) - (1-d)\lambda^{-d} \\ &= (1-d)\left(1 - \frac{1}{\lambda^d}\right) > 0 \quad \forall \lambda > 1, 0 < d < 1 \\ &\Rightarrow f(\lambda) > 0 \quad \forall \lambda > 1 \end{aligned}$$

This demonstrates that the conditions in (A.27) and (A.28) do hold. Hence (i) and (ii) of Proposition 3.

Given $R(1) > 1$, $R(\lambda) < 1$ and $R'(g) < 0$, we can draw the function $R(g)$ for given λ as follows: The function $R(g)$ is illustrated in Figure A1 for $\lambda = 2$ and $\beta = 0.5$. The degree of bias is the distance of R from unity. It is clear that this distance

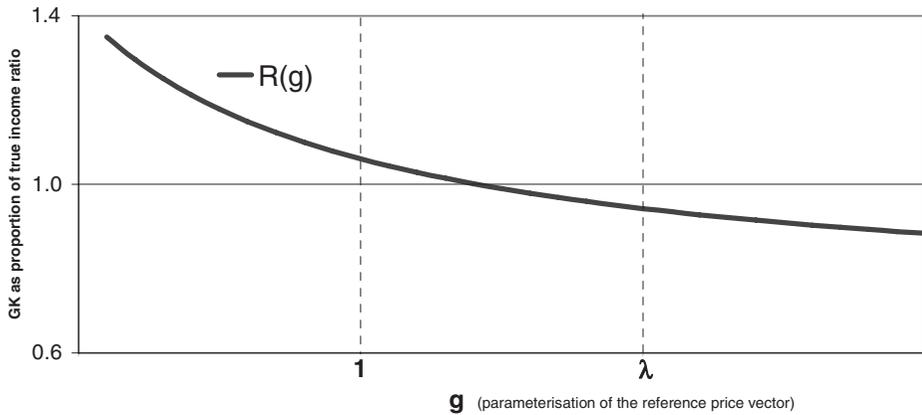


Figure A1. Bias in the GK Income Ratio

Note: The function $R(g)$ is drawn from equation (A.25) for $\lambda = 2$ and $\beta = 0.5$.

increases either as g is larger than λ , or as g becomes smaller than 1. Hence part (iii) of Proposition 3.

Finally, for a given value of g , we can differentiate the expression in (A.25) with respect to λ :

$$(A.31) \quad [\beta g + (1 - \beta)] \frac{\partial R(\lambda; g)}{\partial \lambda} = \frac{\beta(1 - \beta)(\lambda - g)}{\lambda^{1 - \beta}}$$

$$\Rightarrow \frac{\partial R(\lambda; g)}{\partial \lambda} \gtrless 0 \quad \text{as} \quad \lambda \gtrless g$$

If g is greater than λ , as postulated in part (i), R is less than unity and decreasing in λ . Thus the magnitude of the (downward) bias is increasing in λ . If g is less than unity, as postulated in (ii), then λ is greater than g and R is greater than unity and increasing in λ . Thus the magnitude of the (upward) bias is increasing in λ . Hence part (iv) of Proposition 3.

REFERENCES

- Ackland, Robert, Steve Dowrick, and Benoit Freyens, "Measuring Global Poverty: Why PPP Methods Matter," *Annual Conference of the International Association for Research into Income and Wealth*, Cork, Ireland, 2004.
- Afriat, S. N., "The True Index," in A. Ingham and A. M. Ulph (eds), *Demand, Equilibrium, and Trade: Essays in Honor of Ivor F. Pearce*, Macmillan, New York, 37–56, 1984.
- Atkinson, Anthony B. and Andrea Brandolini, "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, 2001.
- Balassa, B., "The Purchasing Power Parity Doctrine: A Reappraisal," *Journal of Political Economy*, 72(6), 584–96, 1964.
- Bergstrand, Jeffrey H., "Structural Determinants of Real Exchange Rates and National Price Levels: Some Empirical Evidence," *American Economic Review*, 81, 325–34, 1991.
- Berry, Albert, François Bourguignon, and Christian Morrisson, "Changes in the World Distribution of Income between 1950 and 1977," *Economic Journal*, 93(37), 331–50, 1983.
- Bhagwati, Jagdish, "Why Are Services Cheaper in the Poor Countries," *Economic Journal*, 94, 279–86, 1984.
- Bourguignon, François and Christian Morrisson "Inequality Among World Citizens: 1820–1992," *American Economic Review*, 92(4), 727–44, 2002.
- Chotikapanich, Duangkamon, "A Comparison of Alternative Functional Forms for the Lorenz Curve," *Economics Letters*, 41(2), 129–38, 1993.
- Chotikapanich, Duangkamon, Rebecca Valenzuela, and D. S. Prasada Rao, "Global and Regional Inequality in the Distribution of Income: Estimation with Limited and Incomplete Data," *Empirical Economics*, 22(4), 533–46, 1997.
- Deininger, Klaus and Lyn Squire, "A New Data Set Measuring Income Inequality," *World Bank Economic Review*, 10(3), 565–91, 1996.
- Diewert, Erwin, "Similarity and Dissimilarity Indexes: An Axiomatic Approach," *University of British Columbia Discussion Paper*, 02–10, 2002.
- Dowrick, Steve and John Quiggin, "True Measures of GDP and Convergence," *American Economic Review*, 87(1), 41–64, 1997.
- Falvey, Rodney E. and Norman Gemmill, "A Formalisation and Test of the Factor Productivity Explanation of International Differences in Service Prices," *International Economic Review*, 37(1), 85–102, 1996.
- Firebaugh, Glenn, "Empirics of World Income Inequality," *American Journal of Sociology*, 104(6), 1597–630, 1999.
- Gerschenkron, Alexander, *A Dollar Index of Soviet Machinery Output, 1927–28 to 1937*, Rand Corporation, Santa Monica, CA, 1951.
- Grosh, Margaret-E. and E. Wayne Nafziger, "The Computation of World Income Distribution," *Economic Development and Cultural Change*, 34(2), 347–59, 1986.

- Hill, Robert J., "Measuring Substitution Bias in International Comparisons Based on Additive Purchasing Power Parity Methods," *European Economic Review*, 44(1), 145–62, 2000.
- Korzeniewicz, Roberto Patricio and Timothy Patrick Moran, "World Economic Trends in the Distribution of Income, 1965–92," *American Journal of Sociology*, 102(4), 1000–39, 1997.
- Kravis, Irving B., Alan Heston, and Robert Summers, *World Product and Income: International Comparisons of Real Gross Products*, The Johns Hopkins University Press, Baltimore, 1982.
- Li, Hongyi, Lyn Squire, and Heng-fu Zou, "Explaining International and Intertemporal Variations in Income Inequality," *Economic Journal*, 108(446), 26–43, 1998.
- Maddison, Angus, *Monitoring the World Economy: 1820–1992*, Organization for Economic Co-operation and Development, Paris and Washington, D.C., 1995.
- Melchior, Arne, Kjetil Telle, and Henrik Wiig, *Globalisation and Inequality: World Income Distribution and Living Standards, 1960–98*. Studies on Foreign Policy Issues, Report 6B, Royal Norwegian Ministry of Foreign Affairs, Oslo, Norway, 1–42, 2000.
- Milanovic, Branko, "True World Income Distribution, 1988 and 1993: First Calculations Based on Household Surveys Alone," *Economic Journal*, 112(476), 51–92, 2002.
- Nuxoll, Daniel A., "Differences in Relative Prices and International Differences in Growth Rates," *American Economic Review*, 84(5), 1423–36, 1994.
- Sala-i-Martin, Xavier, "The World Distribution of Income (Estimated from Individual Country Distributions)," *NBER Working Papers: 8933*, National Bureau of Economic Research, Inc., 2002a.
- , "The Disturbing 'Rise' of Global Income Inequality," *NBER Working Paper*, 8904, 1–72, 2002b.
- Samuelson, Paul, "Theoretical Notes on Trade Problems," *Review of Economics and Statistics*, 46(2), 145–54, 1964.
- , "Analytical Notes on International Real-Income Measures," *Economic Journal*, 84(335), 595–608, 1974.
- Schulz, T. Paul, "Inequality in the Distribution of Personal Income in the World: How it is Changing and Why," *Journal of Population Economics*, 11, 307–44, 1998.
- Summers, Robert and Alan Heston, "The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950–1988," *Quarterly Journal of Economics*, 106(2), 327–68, 1991.
- Theil, Henry, "World Income Inequality," *Economics Letters*, 2(1), 99–102, 1979.
- , *Studies in Global Econometrics*, Kluwer Academic Publishers, Amsterdam, 1996.
- Theil, Henry and James L. Seale, "The Geographic Distribution of World Income, 1950–90," *The Economist*, 4, 1994.
- UNDP, *Human Development Report 1999*, Oxford University Press, New York, 1999.
- Varian, Hal R., "Non-Parametric Tests of Consumer Behaviour," *Review of Economic Studies*, 50(1), 99–110, 1983.
- World Bank, *Purchasing Power of Currencies: Comparing National Incomes Using ICP Data*, International Economics Department, World Bank, Washington, D.C., 1993.
- , "ICP Regional Datasets for 1993," supplied to authors by the Bank's international comparison section, 2000.
- , "Global Development Network Database: Macro Time Series," <http://www.worldbank.org/research/growth/GDNdata.htm>, 2001.