# Making Inferences About The Polarization, Welfare and Poverty Of Nations:
## A Study of 101 Countries 1970-1995

Gordon Anderson

Economics Department
University of Toronto

June 2003

**Introduction.**

The growing gap between rich and poor nations (Jones (1997), (1997a), Kremer et al (2000), Pritchett (1997) and Quah (1997), (2001)) has been a focus of attention in many spheres of economics. For growth theorists it constitutes the challenge to early neoclassical theories of national production that prompted much study of the issue of convergence in Per Capita Gross Domestic Product (Barro(1998))[1]. For those concerned with the plight of the poor it conveys a sense of increasing relative global poverty in terms of Per Capita Gross National Product (Wade (2001)). For those interested in welfare and inequality issues it is of major import since inter-nation differences in income measures account for the largest part of global inequality (Berry et. Al. (1983), Schultz (1998), Bourguignon and Morrison (1999)). Underlying concern about the gap is an implicit "World Welfare Function" relating to the distribution of incomes across nations that ceteris paribus weights the gap as a negative characteristic. Its widening does not imply, and is not implied by, lower global economic welfare, greater inequality or greater absolute poverty but has more to do with Polarization, a relocation of mass from the center towards the tails of the distribution (Esteban and Ray (1994), Foster and Wolfson(1992), Wolfson (1994)).

Polarization has an inter-temporal dimension implying a tendency toward the emergence and increasing intensity and/or separateness of multiple bumps in a distribution over time. As such its investigation has been approached several different ways. Obviously it can occur before bumps appear so that divergent bi-modality (Bianchi (1997)) is not necessary for its existence. Interpreting the global distribution of national per capita incomes as a mixture of rich and poor nation group distributions, Polarization is concerned with how changes in their respective masses

---

[1] For economies starting from different initial disequilibrium points with identical characteristics defining equilibrium national income, Neoclassical growth theory predicts convergence, or less inequality, in the distribution of national incomes over time. However the recent empirical literature investigating the sources of the progress of national per capita incomes is at pains to assert that nations are not identical (Lee, Pesaran and Smith (1997), (1997a)), that the conditional convergence hypothesis underlaying the models does not imply convergence (or more equality) in its distribution per se (Quah(1993), Hart(1995)) and that empirical analysis is better served by models characterized by rich and poor clubs (Quah (1997)).

are reflected in changes in the global distribution. The issue transcends consideration of simple location and scale movements. For example, left skewing poor club and / or right skewing rich club distributions without changing their respective means or variances can engender a widening gap in the global distribution without sub-group location or scale changes. It follows that identifying divergent sub population means and diminishing sub-population variances (Paap and van Dijk (1998)) is not necessary for polarization[2]. The phenomena has also been examined by representing intra distributional dynamics as a Markov Chain (Quah(1997)). In this context period by period mass relocation is viewed as part of an ongoing structurally constant dynamic process wherein polarization is characterized by the relative magnitudes of the constant parameters (transition matrix) of the process. However polarization may be advancing or retarding at different rates over time with consequent advantages to examining the phenomenon period to period in an unconstrained fashion. Finally Esteban and Ray (1994), Wolfson (1994) and Beach and Slotsve (1996) provide indices for ranking the extent of polarization without having to identify the existence or otherwise of multiple modes. Unfortunately, without a distributional theory for the indices, it is not possible to determine if different index values are due to sampling variation or underlaying distributional differences and, like the Gini inequality coefficient, whilst the rankings are complete they may be ambiguous and rank polarized states differently.

Here stochastic dominance techniques[3] are adapted and employed to study the extent of polarization and convergence. While offering only a partial ordering of distributions, they do provide methods for inferentially comparing mass location unambiguously, typically by evaluating various degrees of right separation or left separation between distributions. They can be applied in

---

[2]Bianchi (1997) and Paap and van Dijk (1998) follow a statistics literature which argues that multimodality is more easily studied in the context of mixtures of unimodal distributions (there are relatively few parametrically specified multimodal distributions) proposes tests for spotting multiple modes or dips (Cox (1966), Good and Gaskins (1980), Silverman (1981) and Hartigan and Hartigan (1985)).

[3] See Atkinson (1970) (1987), Kolm (1976), Foster and Shorrocks (1988) for the underlaying theory and Anderson (1996), Davidson and Duclos (2000) and Barrett and Donald (1999) for the statistical implementation.

a period by period fashion without imposing any structure on the nature of the polarization process either in the form of the parameters of the underlying mixture distributions or on the parameters of the Markov Chain transition matrix. Furthermore they may be modified to examine and identify various types of mass relocation from the middle of a distribution toward its tails regardless of whether or not polarization has resulted in the development of multiple modes. Here these techniques are implemented on the per capita GNP of a sample of 101 countries over the years 1970 to 1995. For comparison a corresponding analysis of global welfare and poverty is presented together with values of various polarization and inequality indices.

A rationale for the use of Per Capita GNP measure, as opposed to a Per capita GDP most popular in the study of convergence in the empirical growth literature, is appropriate. The difference, (GNP = GDP plus investment income received from foreign sources less investment income paid to foreign sources) can be substantial[4] within a particular country. GDP measures the output and hence the income produced within a country and is most appropriately employed when the development of the productive capacities of nations is of interest as it is in the empirical growth literature. GNP measures the incomes received within a country and more adequately reflects a countries consumption capacity and hence welfare. Here interest is focused on the welfare of the collective societies in the sample and hence the GNP measure will be used. Section 1 discusses the relationship between mass relocation and the concept of convergence employed in the economic growth literature. Welfare and polarization issues as they relate to stochastic dominance and mass relocation are considered in Section 2. The statistical tests are outlined in Section 3, the results are reported in Section 4 and conclusions are drawn in Section 5. The evidence strongly suggests that polarization, global welfare and poverty took quite different paths.

---

[4] Indeed a companion study currently in progress for a similar collection of countries over a similar time period using per capita GDP data from the Penn World Tables yields the opposite inferences with respect to polarization to those reported here. Whether this is largely a result of the poorer nations being the debtor nations in the sample and thus having their consumption rather than their productive capacities constrained accordingly or whether it is a function of purchasing power parity as opposed to standard exchange rate based comparisons is under ongoing investigation.

# 1. Convergence and Mass Relocation.

As Bernard and Durlauf (1996) observe, the study of convergence in the empirical growth literature has followed two routes. One, interpreting convergence as catching up, thinks of two economies i and j converging in terms of $E(abs(y_{i,t+T} - y_{j,t+T}) \mid \xi_t) < abs(y_{it} - y_{jt})$ for some $T > 0$, where $Y_{it}$ is per capita income in country i in period t, $y_{it} = \ln(Y_{it})$ and $\xi_t$ is information at time t. The other, interpreting convergence as equality of long term forecasts, contemplates $\lim_{T \to \infty} E(abs(y_{i,t+T} - y_{j,t+T}) \mid \xi_t) = 0$ as the convergence condition. When i and j have stationary and structurally identical growth processes both definitions of convergence are satisfied. However one can configure stationary growth processes for i an j with different structures that violate both definitions and further, if i and j have identical non-stationary structures that are not co-integrated, the second definition will be violated.

The stationary - non stationary distinction has antecedents. Gibrat (1930) whose work provided theoretical foundation for using log normality in income distribution analysis, formulated a growth process for $Y_{it}$ starting with the premise that the initial value of the variate $Y_{i0}$ is subject to a sequence of mutually independent proportionate changes $e_{ik}$, k=1,..,t so that after the passage of time t, $Y_{it} = Y_{i0}(1+e_{i1})(1+e_{i2}) ...(1+e_{it})$. Assume $|e_{ik}|$ to be small relative to 1 and let $\ln(1+e_{ik}) = u_{ik}$ and $y_{it} = \ln Y_{it}$ where $u_{ik}$ is an i.i.d. process with $E(u_{it})=\mu_i$ and variance $\sigma^2$ then:

$$y_{it} = \mu_i + y_{it-1} + u_{it} \qquad [1]$$

with $\mu_i$ (again small relative to one) corresponding to the incremental drift or growth in the process and $u_{it}$ corresponding to the increment of a drifting Weiner process. Gibrat demonstrated that, after sufficient passage of time t, this renders $N(\ln Y_{i0}+(\mu_i-\sigma/2)t,\sigma^2 t)$ the distribution of $y_{it}$. Generally this process would not satisfy either of the Bernard and Durlauf conditions for $i \neq j$, even if they had the same parametric structure, unless the processes were co-integrated. Kalecki (1945) proposed an alternative process which, in the present context, replaces [1] with:

$$y_{it} = \eta_i + \mu_i t + (1-\lambda_i)y_{it-1} + u_{it} \qquad [1a]$$

where $1 > \lambda_i > 0$. Kalecki establishes that, after a sufficient passage of time, the distribution of $y_{it}$ will be $N((\eta_i+\mu_i t)/\lambda_i, \sigma^2/\lambda_i^2)$. Again $\mu$ may be construed as the incremental growth component but in this case the logarithm of the proportionate change in Y is negatively related to lnY via $-\lambda$, note also the variance of this process is constant through time and the distribution is independent of the initial starting value. Clearly this process would satisfy the Bernard and Durlauf conditions for $i \neq j$ as long as they had the same parametric structure. Unlike Gibrat's model, [1a] may be re-written as a partial adjustment model with $(\eta+\mu t)/\lambda$ as the target and $\lambda$ as the adjustment rate. In case [1] the distribution of y is divergent through time and in case [1a] the distribution may be thought of as convergent or at least non-divergent in the sense that whilst both models predict increasing means [1] predicts increasing dispersion whereas [1a] does not.

When the processes i and j have different parametric structures the distribution of $y_t$ becomes a mixture in both cases. Differing parametric structures may arise when different processes are associated with groups with different characteristics. In the present context economic models of the separation of rich and poor groups abound (the threshold model of Azariadis and Drazen (1990) employed in Durlauf and Johnson (1995), the club model of Galor and Zeira (1993) employed in Quah (1997)). As Durlauf and Quah (2002) point out, regression models essentially consider conditional averages and are uninformative as to whether the poor - rich gap is closing which requires study of the complete distribution over time. This is even more pertinent when the progress of welfare, poverty and polarization is at issue since it is ultimately depends upon the nature of mass relocation which is not completely reflected in movements of conditional averages.

## 2. Welfare, Poverty and Stochastic Dominance: Thinking About Mass Relocation.

Atkinson (1970), Kolm (1976) and Foster and Shorrocks (1988) highlight the importance of the nature of mass relocation in income distributions for empirical welfare comparisons. They provide specific definitions of the distributional change necessary and sufficient to engender a welfare improvement for welfare functions in particular classes. The change is defined in terms of Stochastic Dominance Orderings which emerge from considering the average utility gained in

moving from one income distribution to another. Consider $\delta$, the change in the expected value of societal utility $u(x)$[5] which has the properties $(-1)^{j-1}\partial^j u/\partial x^j \geq 0$ $j = 1,..,i$ for some $i > 0$, based upon moving from density function $G(x)$ to $F(x)$ both defined on the interval [a,b], it may be written as:

$$\delta = E_F(u(x)) - E_G(u(x)) = \int_a^b u(x)(dF - dG)$$

A necessary and sufficient condition for $\delta > 0$ for a given $i$ is:

$$\int_a^x (F_{i-1}(z) - G_{i-1}(z))dz \leq 0 \text{ for all } x \wedge \int_a^x (F_{i-1}(z) - G_{i-1}(z))dz < 0 \text{ for some } x \in [a,b] \quad [2]$$

where, letting $f(x) = F_0(x)$, $F_i(x)$ is defined recursively as:

$$F_i(x) = \int_a^x F_{i-1}(z)dz: \quad (x \leq b, i \geq 1)$$

---

[5] Generally $E(u(x))$ is thought of as the welfare function but if $u(x) = -P(x)$ where $P(x)$ is a poverty index based upon incomes, the same dominance criteria can be used to evaluate poverty states measured by poverty indices in a given class (Atkinson (1987)). In terms of social welfare First Order Dominance corresponds to an ordering of social preferences based upon monotonic utilitarian social welfare functions, Second Order to a social preference for mean preserving progressive transfers and Third Order to a social preference for mean preserving progressive transfers at lower income levels. In the context of poverty indices, different levels of dominance ensure for any poverty line the same direction of change for all indices in the class defined by the order of dominance. Hence First Order Dominance implies coherence between all continuous non-decreasing in X poverty measures (e.g. poverty counts), Second Order implies coherence between all continuous non-decreasing weakly concave in X poverty measures (e.g. the average deviation from the poverty line of the poor) and Third Order implies coherence between all continuous non-decreasing strictly concave measures (e.g. the average squared deviation from the poverty line of the poor). Atkinson (1987) presents visual representations of measures from these various classes.

and $G_i(x)$ is defined similarly. When [2] is satisfied $f(x)$ is said to stochastically dominate $g(x)$ at order i. In the following $f(x) \succeq_j g(x)$ denotes dominance of $g(x)$ by $f(x)$ of at least order j. For convenience $f(x) \succ_j g(x)$ denotes strict order j dominance where strict inequality in [2] obtains over the relevant range. Note for $i < j$, $f(x) \succeq_i g(x)$ implies $f(x) \succeq_j g(x)$, furthermore the relationship is transitive in that if $f(x) \succeq_j g(x)$ and $g(x) \succeq_j h(x)$ then $f(x) \succeq_j h(x)$. Though the ordering is not complete it is unambiguous and, given the properties of u( ), facilitates orderings of unobservable distributions of u(x) in terms of observable distributions of x.

Here it is convenient to interpret i'th order dominance as the degree of "i'th Order Right Separation" of the two distributions. When $f(x) \succeq_i g(x)$, $F_i(x_1) = G_i(x_2)$ implies $x_2 \leq x_1$, so that $F_i$ is everywhere not to the left of $G_i$ and to the right of it at least somewhere, implying a sense of Right Separation of $f(x)$ from $g(x)$ at the i'th level of integration. As limiting examples let x be a transformation of y with respective distribution functions $g(y)$ and $f(x)$, then a positive location shift transformation[6] implies $f(x) \succeq_1 g(x)$, if the transformation is a location preserving, scale reducing shift then $f(x) \succeq_2 g(x)$ and if the transformation is a location and scale preserving, positive skewing shift then $f(x) \succeq_3 g(x)$.

Of equal of interest is the idea of "i'th Order Left Separation" characterized by a condition of the form:

$$\int_x^b (F_{i-1}(z) - G_{i-1}(z))dz \leq 0 \textit{ for all } x \wedge \int_x^b (F_{i-1}(z) - G_{i-1}(z))dz < 0 \textit{ for some } x \quad [2]$$

----

[6]This is easily demonstrated for distributions confined to the positive orthant, $f(x) \succeq_1 g(x)$ is sufficient for $E(x|f) > E(x|g)$ since:

$$E(x|f) - E(x|g) = \int_0^\infty [(1-F(x))-(1-G(x))]dx > 0 \rightarrow \int_0^\infty (G(x)-F(x))dx > 0$$

Defining w = -x and f(w) and g(w) as appropriately transformed distributions on [-b,-a] this condition is equivalent to $f(w) \succeq_i g(w)$ (with $F_i(w)$ and $G_i(w)$ defined as before) and has the analogous " 'i'th order Left Separation of f(x) from g(x)" interpretation[7]. In this context the relationship $f(w) \succeq_1 g(w)$ may be thought of as a negative location shift transformation, if the transformation is a location preserving scale reducing shift then $f(w) \succeq_2 g(w)$ and if the transformation is a location and scale preserving, negative skewing shift then $f(w) \succeq_3 g(w)$.

Assuming relative club sizes remain constant, polarization between rich and poor countries may now be thought of in terms of the rich club distribution Right Separating and the poor club distribution Left Separating at some order (of course one club separating in the appropriate direction whilst the other remains unchanged would also constitute polarization). When the club distributions are separately identified, polarization can be examined statistically by performing the relevant stochastic dominance tests jointly on successive realizations of the relevant club distributions. Letting $f^j(x)$ and $g^j(x)$ be the period 'j' rich club and poor club distributions respectively, three conditions need to hold simultaneously for i'th order polarization:

1) $f^1(x) \succeq_1 g^1(x)$ (establishing that the rich club is first order right separated from the poor club).
2) $f^2(x) \succeq_i f^1(x)$ (establishing that the rich club at least i'th order right separates in period 2).
3) $g^2(w) \succeq_i g^1(w)$ (establishing that the poor club at least i'th order left separates in period 2).

Thus as limiting cases First Order Polarization is engendered by the respective club means moving further apart, Second Order Polarization arrises when the clubs become more concentrated around their respectively unchanged means and Third Order Polarization occurs when the poor club skews left and the rich club skews right with their means and variances remaining unchanged.

When, as in the present case, the observed distribution is an unknown mixture of unobserved rich and poor country distributions, the problem is to analyse the consequences of polarization

---

[7] This type of dominance is used in the finance literature and relates to the analysis of risk loving behaviour (see Levy and Weiner (1998)).

within the observed mixtures. Inferences can be made by associating the lower and upper tails of the observed mixtures with the respective poor and rich clubs. Thus partitioning the distributions at some common defining point x* (in the present case it will be the pooled sample medians respectively) and considering the relative progress of the distributions $f^1(x|x<x^*)$, $f^2(x|x<x^*)$, $f^1(x|x>x^*)$ and $f^2(x|x>x^*)$ two conditions need to hold simultaneously:

1m) $f^2(w|x<x^*) \succeq_i f^1(w|x<x^*)$ (the left tail at least i'th order left separates in period 2).

2m) $f^2(x|x>x^*) \succeq_i f^1(x|x>x^*)$ (the right tail at least i'th order right separates in period 2).

Clearly $f^1(x|x<x^*) \preceq_1 f^1(x|x>x^*)$ is always true in this case and does not need to be established so that an analogue to condition 1) employed when both rich club and poor club are separately observed is no longer required.

## 3. Statistical Tests.

Tests for mass relocation (stochastic dominance) conditions have proliferated in the literature in recent years, Anderson (1996) employs the distribution of integral approximations, Davidson and Duclos (2000) employ the distribution of incomplete moments and McFadden (1989) and Barrett and Donald (1999) employ distributions of functions of the empirical distribution function. The first two families of tests are attractive because they are easily adapted to situations where samples are non i.i.d. (Anderson (1998), (2003), Davidson and Duclos (2000)). Essentially they are a sequence of joint inequality tests for examining $v_i(f,g)$ a vector of asymptotically normally distributed estimates of $F_i(x)-G_i(x)$ at a selection of pre-specified values of x. The latter approach is attractive because, unlike the first two approaches, it is a consistent test focussing on the maximum distance $F_i(x)-G_i(x)$ over the whole range of x, however it can be shown that, under smoothness assumptions, the inconsistency problem is not substantive (Anderson (2000)). Anderson (2001) provides a taxonomy of tests appropriate for examining both within and

between population polarization[8]. Here the tests formulated in Anderson (2001) modified to account for the between sample dependence engendered by the "panel" type nature of the data as outlined in Anderson (2003) are used.

The practice has been to employ either the Maximum Modulus Distribution (tables for which are provided in Stoline and Ury (1978)) to a collection of asymptotically standard normal statistics (which is a conservative test) based upon the K individual elements of the vector $v_i(f,g)$ and their corresponding standard deviations, or to employ the joint testing procedures advocated in Kodde and Palm (1986) and Wolak (1989). The advantage of the former is that the inequality relations may be studied in detail, the advantage of the latter, it is not a conservative test. For the joint test let $v_{iw}(f,g)$ be the inequality constrained estimate of the vector $v_i(f,g)$ and let $(\Omega_{vi})^+$ be a generalized inverse of the covariance matrix of $v_i$, then for:

$$W = (v_i(f,g) - v_{iw}(f,g))'(\Omega_{v_i})^+(v_i(f,g) - v_{iw}(f,g))$$

the distribution of W is such that:

$$P(W \geq c) = \sum_{i=0}^{k-1} P(\chi^2_i \geq c)w(k,k-i,\Omega_{v_i})$$

where $w(k,k-i,\Omega)$ is a weight function corresponding to the probability that v, with covariance

---

[8] A cautionary note is in order. These tests are designed to detect the "hollowing out of" the centre of the distribution (Beach, Chaykowski and Slotsve (1998)) identified with within distribution polarization. Unfortunately it can be shown that when the distribution to hand is a mixture of two closely located sub distributions and polarization takes the form of limited reductions of sub population variances (increased concentration around the respective poles), polarization will manifest itself in the observed mixture as an **increase** in central mass. Fortunately it appears that this phenomena only occurs when the sub population distributions are located fairly close together and long before bi-modality occurs in the mixture (which, for example in a 50/50 mixture of normals with equal variances less than one, occurs when the means are more than one standard deviation apart). Thus we can be sure that, if the mixture is bi-modal, sub population polarization will manifest itself in a loss of mass at the centre of the distribution.

matrix $\Omega$, has k-i-1 of its k-1 independent elements positive. Closed form expressions for the weight function only exist for k-1 up to 4, however following the suggestion in Wolak (1989) they can readily be approximated via pseudo normal random number generation.

For the polarization tests the vector $v_i(f,g)$ is redefined as:

$$v_i(f^1,f^2) = \begin{array}{c} v_i(f^1(w|x<x^*), f^2(w|x<x^*)) \\ v_i(f^1(x|x>x^*), f^2(x|x>x^*)) \end{array}$$

with the covariance matrix redefined accordingly.

## 4. Results.

Data from the World Bank World Development Indicator series on per capita GNP in 1987 $US constant prices and population size were collected for 101 countries for the years 1970, 1978, 1987 and 1995. The countries, listed in Appendix 1, were selected on the basis of having complete series for the period, the most notable omissions being the USSR, Hungary, Poland, China, East and West Germany and Austria. The years were selected as roughly equal spaced intervals over the observation period. Population size was collected for the purpose of sample weighting, which is an issue for two reasons. Firstly when using these techniques to study individual welfare with household based data, household size re-weighting is crucial for theoretical consistency (though practically it rarely seems to affect the qualitative nature of the results because of the limited variation in household size). Here the nation is the household and, with much greater variation in its size, re-weighting sample observations by relative population size is potentially more important from an individual welfare perspective. Secondly, the testing and Kernel Estimation techniques employed assume within year i.i.d. sampling and re-weighting is necessary to undo the stratified sampling inherent in the data. Results on both weighted and un-weighted samples are presented for comparison. Similarly the use of a panel abrogates the i.i.d. assumptions usually invoked in this work. Again for comparison, results under an assumed i.i.d.

12

and a panel based dependent sample scheme are reported.

Table 1 presents the summary statistics for log per capita GNP. Notable from the table is the upward shift in location of the underlaying distribution that is somewhat less obvious in the weighted sample. The substantially increased weight that poor countries have in the weighted sample suggests that mean incomes were increasing more in the rich country club than in the poor country club. The spread on the other hand has steadily increased throughout the period, though the rate is again greater in the unweighted sample suggesting relatively greater divergence in the rich as opposed to poor country club. In terms of location 1970-1978 constituted a big leap forward whilst there was a fall back in the '78 - '87 period and some slight recovery in the '87-'95 period. Both weighted and unweighted distributions continued to spread throughout the period which is also reflected in a continuing increase in the range of the distribution. To put this into perspective the ratio of the richest to the poorest country per capita incomes was approximately 83 in 1970 and 343 in 1996![9]. What is not obvious from the table is the multi-modal nature of the

| Table 1. Statistics for the Natural Logarithm of Per Capita Gross National Product (constant $US) weighted and unweighted by population share. | | | | |
|---|---|---|---|---|
| | 1970 | 1978 | 1987 | 1995 |
| Mean | 7.1220374 | 7.5156034 | 7.3771270 | 7.3948625 |
| Weighted Mean | 7.0276138 | 7.2673949 | 7.1216338 | 7.0803898 |
| Median | 6.8721650 | 7.3263788 | 7.1065725 | 7.2262794 |
| Weighted Median | 6.5238583 | 6.7694766 | 6.3801225 | 6.5366635 |
| Standard Deviation | 1.1792535 | 1.3092340 | 1.4519676 | 1.6822738 |
| Weighted Std. Dev. | 1.4123980 | 1.5145033 | 1.5626208 | 1.7494431 |
| Maximum | 9.5523804 | 10.147129 | 10.037144 | 10.383178 |
| Minimum | 5.1375640 | 5.1970800 | 5.0751738 | 4.5442333 |

[9] Over the four sample years, the United States, Kuwait, Switzerland and Luxembourg were the richest and Malawi, Bangladesh, Malawi and Zambia the poorest nations respectively.

distributions reflected in the plots of the Kernel density estimates[10] in diagrams 1 and 2 corresponding to unweighted and weighted samples respectively. The distributions are essentially bi-modal (alleviating aforementioned concerns regarding sub-population polarization engendering an increased central mass of the mixture) and more obviously so in the weighted sample case.

**Table 2. Inequality and Polarization Indices for the Natural Logarithm of Per Capita Gross National Product (constant $US).**

|  | 1970 | 1978 | 1987 | 1995 |
|---|---|---|---|---|
| Unweighted Gini of lnGNP [1] (inequality) | 0.0932 | 0.0985 | 0.1118 | 0.1296 |
| Weighted Gini of lnGNP [1,2] (inequality) | 0.1096 | 0.1128 | 0.1152 | 0.1302 |
| Inter-quartile Range / $\sigma$ (polarization) | 1.4043 | 1.6496 | 1.5952 | 1.7145 |
| Inter-quartile Range/ Range (polarization) | 0.3744 | 0.4345 | 0.4663 | 0.4936 |
| Esteban - Ray Index ($\alpha$=1)[3] (polarization) | 0.0138 | 0.0156 | 0.0148 | 0.0172 |

**Table Notes.**
[1] Strictly speaking this is not a Gini coefficient since it relates to the logarithm of income rather than its level.
2. A sample weighted Gini of the logarithms of incomes is equivalent to The Esteban - Ray Polarization index with $\alpha = 0$.
3. A range of values for $\alpha$ (.5, 1, 1.5) each yielded the same qualitative direction of the Polarization index.

The extent to which this bi-modality feature is reflected in various Polarization indices is reported in Table 2. Note that while the weighted and unweighted Gini coefficients indicate a monotonic increase in inequality throughout the sample period, the Esteban- Ray[11] and inter-

---

[10]An Epanechnikov kernel (Silverman (1986)) was employed in each case.

[11]The Esteban and Ray (1994) Polarization Index for the discretely distributed random variable y which takes on any one of n values $y_i$ with probabilities $\pi_i$, i=1,..,n. For constants K and $\alpha$, is of the form:

$$P(\pi, y) = K \sum_{i=1}^{n} \sum_{j=1}^{n} \pi_i^{1+\alpha} \pi_j |y_i - y_j|$$

K is a multiplicative constant which does not affect the ordering, $\alpha$ is a parameter reflecting the polarization sensitivity of the index where $0 < \alpha \leq 1.6$, the larger its value the further is the

14

quartile range/standard deviation polarization indices record an increase, decrease and then an increase in polarization whilst the inter-quartile range/range polarization index indicates a monotonic increase in polarization.

Table 3 reports social welfare comparisons based upon stochastic dominance criteria. The results clearly indicate that, regardless of assumed sampling model or population weighting

**Table 3. Stochastic Dominance Rankings of Per Capita Gross National Product Distributions[1,2,3].**

| Comparison Years | iid unweighted | iid weighted | dependent smpl unweighted | dependent smpl weighted |
|---|---|---|---|---|
| 1970-1978 | (↑, 1) [1.000, 0.020] | (↑, 1) [0.302, 0.000] | (↑, 1) [0.998, 0.000] | (↑, 1) [0.969, 0.000] |
| 1978-1987 | (↓, 2) [0.186, 0.968] {0.010, 0.535} | no decision [0.941, 0.696] {0.536, 0.130} | (↓, 1) [0.000, 0.365] | (↓, 1) [0.023, 0.908] |
| 1987-1995 | (↓, 2) [0.308, 0.753] {0.012, 0.541} | (↓, 1) [0.000, 0.972] | (↓, 1) [0.021, 0.075] | (↓, 2) [0.180, 0.533] {0.005, 0.611} |
| 1970-1995 | no decision [0.001, 0.014] | (↓, 1) [0.000, 0.393] | no decision [0.015, 0.000] | (↓, 2) [0.217, 0.014] {0.000, 0.627} |

**Notes:**
**1. (↑, i) indicates a social welfare improvement of order "i" and (↓, i) indicates a social welfare decline of order "i" based upon a P(Null) < .05 decision criteria.**
**2. $[p_1, p_2]$ Correspond to respective upper tail probabilities of Wald Criteria for the 1st Order Dominance comparison year B dominates year A and year A dominates Year B.**
**3. $\{p_1, p_2\}$ Correspond to respective upper tail probabilities of Wald Criteria for the 2nd Order Dominance comparison year B dominates year A and year A dominates Year B.**

scheme, welfare improved and then diminished over the period. Furthermore in the population weighted i.i.d. and dependent sample versions it did so to the extent that a welfare loss could be

---

measures departure from an inequality measure. Here the i'th countries population share in the overall sample corresponds to $\pi_i$. Note that this index is designed to pick up clustering around many modes whereas the much simpler range based indices are focussed upon bimodal structures and are thus potentially more powerful in the present context.

determined over the whole period. 1970-1978 was unambiguously a period of improvement and 1978-1995 was unambiguously a period of deterioration given the transitivity of stochastic dominance relationships. Neither the weighting scheme nor the statistical model specification appears to have had a substantive effect on the qualitative nature of these results which also have profound implications for a commentary on poverty. Following Atkinson (1987) whatever fixed poverty line is chosen, any measure of absolute poverty based upon a monotonic function of per capita GNP outcomes below that line, would have recorded a reduction in poverty over the 1970-1978 period. The increase in poverty over the 1978-1995 period is equally unambiguous. Note that if poverty or the plight of the poor is viewed as a relative concept, then the degree of polarization is perhaps a more relevant criteria.

| Table 4. Polarization Rankings of Per Capita Gross National Product Distributions[1,2,3,4]. | | | | |
|---|---|---|---|---|
| Comparison Years | iid unweighted | iid weighted | dependent smple unweighted | dependent smple weighted |
| 1970-1987 | no decision [0.836, 0.473] {0.000, 0.011} | ($\uparrow$, 1) [0.797, 0.000] | ($\uparrow$, 1) [1.000, 0.000] | ($\uparrow$, 1) [1.000, 0.000] |
| 1978-1987 | ($\uparrow$, 1) [1.000, 0.006] | ($\uparrow$, 1) [0.999, 0.031] | ($\uparrow$, 1) [0.149, 0.049] | ($\uparrow$, 1) [0.888, 0.000] |
| 1987-1995 | no decision [0.266, 0.971] {0.174, 0.909} | ($\uparrow$, 2) [0.959, 0.079] {0.906, 0.019} | ($\uparrow$, 1) [0.999, 0.000] | ($\uparrow$, 1) [1.000, 0.000] |
| 1970-1995 | ($\uparrow$, 1) [0.999, 0.001] | ($\uparrow$, 1) [0.896, 0.008] | ($\uparrow$, 1) [1.000, 0.000] | ($\uparrow$, 1) [1.000, 0.000] |

Notes:
1. ($\uparrow$, i) indicates polarization of order "i" and ($\downarrow$, i) indicates de-polarization of order "i" based upon a P(Null) < .05 decision criteria.
2. [$p_1, p_2$] Correspond to respective upper tail probabilities of Wald Criteria for the 1st polarization comparison year B relative to year A and year A relative to year B.
3. {$p_1, p_2$} Correspond to respective upper tail probabilities of Wald Criteria for the 2nd Order polarization comparison year B relative to year A and year A relative to year B.
4. For the purposes of the Polarization test the distributions in each of the comparison years were partitioned at the median of the pooled sample.

From Table 4 it is evident that the convolution of inappropriately assuming the data to be

drawn from an i.i.d. scheme and not weighting the data by population size yields results at odds with all other combinations of assumptions which themselves yield a reasonably homogenous body of evidence. This is perhaps not surprising given the impact of re-weighting apparent in Diagrams 1 and 2 and the obvious dependence of successive samples. Polarization between rich and poor countries continued unabated throughout the period and, with the exception of the i.i.d. unweighted results, the evidence is that the trend was steady in each of the sub-periods. The only polarization index in Table 2 consistent with this is the Inter-quartile Range-Range ratio. The sustained polarization may be observed in diagrams 1 and 2 by the ever increasing lateral distance between the two primary points of modality in successive distributions. However a qualification is in order here, the rich club seems to have become relatively smaller in the 1978-87 transition. The rich and poor clubs continue to separate but it appears that their relative sizes have changed which violates the constant relative club size assumption invoked in developing the polarization tests and may well be the source of contradiction between the polarization tests and the polarization indices reported in Table 2. Interestingly the club polarization process continues through periods of welfare improvement and diminishing poverty as well as through periods of welfare deterioration and increasing poverty indicating clearly the possibility of simultaneous polarization and welfare improvement (or absolute poverty reduction). The gap between rich and poor countries grew unequivocally throughout the sample period.

**Conclusions.**

Interpreting convergence (depolarization) and welfare improvement as having to do with the relocation of the mass within a distribution in a particular fashion requires empirical techniques which facilitate assessment of the manner in which mass has relocated. Such techniques for identifying polarization in a collection or mixture of rich and poor countries have been outlined which draw on, and provide companions to, extant stochastic dominance techniques for analysing the progress of global economic well-being and poverty. They do not rely upon the existence of bi-modality in the distribution and do not impose any structure on the polarization process. Employing these techniques in an analysis of the distribution of per capita GNP (representing the

17

consumption capacity of a country) over the period 1970 - 1995 for a broad sample of countries has revealed that, whilst welfare increased and then diminished and poverty diminished and then increased, polarization between rich and poor countries has continued unabated throughout the period.

Re-weighting the data by the relative population size of the country was entertained for both theoretical (attaching the same weight to individuals in different countries) and statistical (undoing the stratified sampling characteristic of the data) reasons. Using unweighted data may be construed as employing a social welfare function over nation states whereas employing weighted data may be thought of as employing a social welfare function over the individuals in those nation states[12]. The question is, how badly is one likely to be misled by using one approach rather than the other? Qualitatively there appears to be little difference in the two sets of results in that there were no ranking reversals though occasionally one approach yielded a "no decision" whereas the other was decisive. Re-weighting appeared to sharpen the polarization and welfare conclusions as did accommodating between observation period dependencies due to the panel-type nature of the data. Perhaps the most significant effect was on the Kernel estimates of the distribution of per capita GNP. The profound differences between the weighted and un-weighted distributions highlight the polarization phenomena confirmed by the statistical tests.

In particular the results emphasize the distinction between polarization and inequality and the notion that one does not imply the other. Thus in a period (1970-1978) when the plight of poor countries improved in terms of their per capita GNP the gap between them and the wealthier nations widened, sustaining the view that the position of the poor worsened in a relative sense. The findings are robust to whether the sample is weighted by population size or not and to whether due allowance is mode for the "panel type" nature of the data.

---

[12]I'm grateful to a referee for suggesting this interpretation.

18

**References.**

Anderson G.J., (1996) "Nonparametric Tests for Stochastic Dominance in Income Distributions" Econometrica 64 pp.1183-1193.

Anderson G.J., (1998) "Liquidity Preference ,Yield Spreads and Stochastic Dominance"mimeo Economics Department University of Toronto.

Anderson G.J. (2000) "A Note on the Consistency of Tests Employing Point-Wise Comparisons for Examining the Equality of Two Functions" mimeo Economics Department University of Toronto.

Anderson G.J. (2001) "Toward an Empirical Analysis of Polarization" mimeo Economics Department University of Toronto.

Anderson G.J. (2003) "Poverty in America 1970-1990: Who Did Gain Ground? An Application of Stochastic Dominance Criteria Employing Simultaneous Inequality Tests in a Partial Panel." Forthcoming Journal of Applied Econometrics.

Atkinson A.B. (1970) "On the Measurement of Inequality" Journal of Economic Theory 2, 244-263.

Atkinson A.B. (1987) "On the Measurement of Poverty" Econometrica 55, 749-764.

Barrett G. and S. Donald (1999) "Consistent Tests for Stochastic Dominance" mimeo Boston University Economics Department.

Barro R.J. (1998) Determinants of Economic Growth: A Cross Country Empirical Study MIT Press.

Beach C.M., R.P. Chaykowski and G.A. Slotsve (1998) Inequality and Polarization of Male earnings in the US 1968-1992 North American Journal of Economics and Finance 8 135-151

Beach C.M.and G.A. Slotsve (1996) "Are we becoming 2 Societies? Income, Polarization and the Middle Class in Canada" C.D. Howe Institute.

Berry A., F. Bourguignon and C. Morrisson (1983) "Changes in the World Distribution of Income Between 1950 and 1977" Economic Journal 93 331-350.

Bianchi M. (1997) "Testing for Convergence: Evidence from Non-Parametric Multimodality Tests" Journal of Econometrics 12 393-409.

Bourguignon F. and C. Morrison (1999) "The Size Distribution of Income Among World Citizens 1820-1990" Mimeo.

Cox D.R. (1966) "Notes on the Analysis of Mixed Frequency Distributions" British Journal of Mathematical Statistics and Psychology 19 39-47.

Davidson, R and J-Y Duclos (2000) "Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality. Econometrica 68 1435-1464.

Esteban, J-M and D Ray (1994) "On the Measurement of Polarization" Econometrica 62, 819-851.

Foster J.E. and A.F. Shorrocks (1988) "Poverty Orderings" Econometrica 56 pp.173-177.

Foster J.E. and Wolfson M.C. (1992) "Polarization and the Decline of the Middle Class: Canada and the U.S." Mimeo Vanderbilt University.

Good I.J. and Gaskins R.A. (1980) "Density Estimation and Bump-Hunting by the Penalised Likelihood Method Exemplified by Scattering and Meteorite Data." Journal of the American Statistical Association 75 42-56.

Hart P.E. (1995) "Galtonian Regression Across Countries and the Convergence of Productivity" Oxford Bulletin of Economics and Statistics 57 287-293.

Jones C.I. (1997) "Convergence Revisited" Journal of Economic Growth 2 131-153.

Jones C.I. (1997a) "On the Evolution of the World Income Distribution" Journal of Economic Perspectives 11 19-36.

Kolm S-C. (1976) "Unequal Inequalities I, II" Journal of Economic Theory 12 416-442, 13 82-111.

Lee K., Pesaran M.H. and R. Smith (1997a) "Growth and Convergence in a Multi-Country Empirical Stochastic Solow Model" Journal of Applied Econometrics 12 357-392.

Lee K., Pesaran M.H. and R. Smith (1997a) "Growth Empirics: A Panel Data Approach. A Comment" Quarterly Journal of Economics 319-323.

Kremer M., A. Onatski and J. Stock (2000) "Searching for Prosperity" Mimeo Economics Department Harvard University.

McFadden D. (1989) "Testing for Stochastic Dominance" in Studies in the Economics of Uncertainty: In Honor of Josef Hadar. Fomby T.K. and Seo T.K. eds. Springer. 113-134.

Paap R. and van Dijk (1998) "Distribution and Mobility of Wealth of Nations" European Economic Review 42 1269-1293.

Pritchett L. (1997) "Divergence, Big Time" Journal of Economic Perspectives 11 3-17.

Quah D.T. (1993) "Galton's Fallacy and Tests of the Convergence Hypothesis" Scandinavian Journal of Economics 95 427-443.

Quah D.T. (1997) "Empirics for Growth and Distribution: Stratification, Polarization, and Convergence Clubs" Journal of Economic Growth 2 27-59.

Quah D.T (2001) "Discussion of "Searching for Prosperity" by Michael Kremer, Alexei Onetski and James Stock" mimeo Economics Department London School of Economics

Schultz T.F. (1998) "Inequality in the Distribution of Personal Income in the World: How is it Changing and Why" Journal of Population Economics 11.

Sen A. (1997) On Economic Inequality (Expanded edition with annexe by J.E. Foster and A. Sen) Clarendon Press Oxford.

Stoline M.R. and H.A. Ury (1979) "Tables of the Studentised Maximum Modulus Distribution and an Application to Multiple Comparisons Among Means" Technometrics 21. 87-93

Wade R. (2001) "Winners and Losers" Economist April 26th.

Wolak F.A. (1989) "Testing Inequality Constraints in Linear Econometric Models" Journal of Econometrics 41 205-235.

Wolfson, M.C. (1994) "When Inequalities Diverge" American Economic Review Papers and Proceedings 84, 353-358.

## Appendix 1. Countries included in the sample.

Algeria, Argentina, Australia, Bahamas, Bangladesh, Barbados, Belgium, Benin, Bolivia, Botswana, Brazil, Burkina Faso, Burundi, Cameroon, Canada, Central African Republic of Chad, Chile, Colombia, Congo, Costa Rica, Cote d'Ivoire, Denmark, Dominica, Dominican Republic, Ecuador, Egypt, El Salvador, Fiji, Finland, France, Gabon, Gambia, Ghana, Greece, Guatemala, Guyana, Haiti, Honduras, Hong Kong, Iceland, India, Indonesia, Ireland, Israel, Italy, Jamaica, Japan, Kenya, Korea, Kuwait, Lesotho, Luxembourg, Madagascar, Malawi, Malaysia, Mali, Mauritania, Mexico, Morocco, Nepal, Netherlands, New Zealand, Nicaragua, Niger, Nigeria, Norway, Oman, Pakistan, Panama, Papua New Guinea, Paraguay, Peru, Philippines, Portugal, Rwanda, Senegal, Seychelles, Sierra Leone, Singapore, South Africa, Spain, Sri Lanka, St. Vincent and the Grenadines, Suriname, Swaziland, Sweden, Switzerland, Syria, Thailand, Togo, Trinidad and Tobago, Tunisia, Turkey, United Kingdom, United States, Uruguay, Venezuela, Zaire, Zambia, Zimbabwe.

Distribution of Per Capita GNP
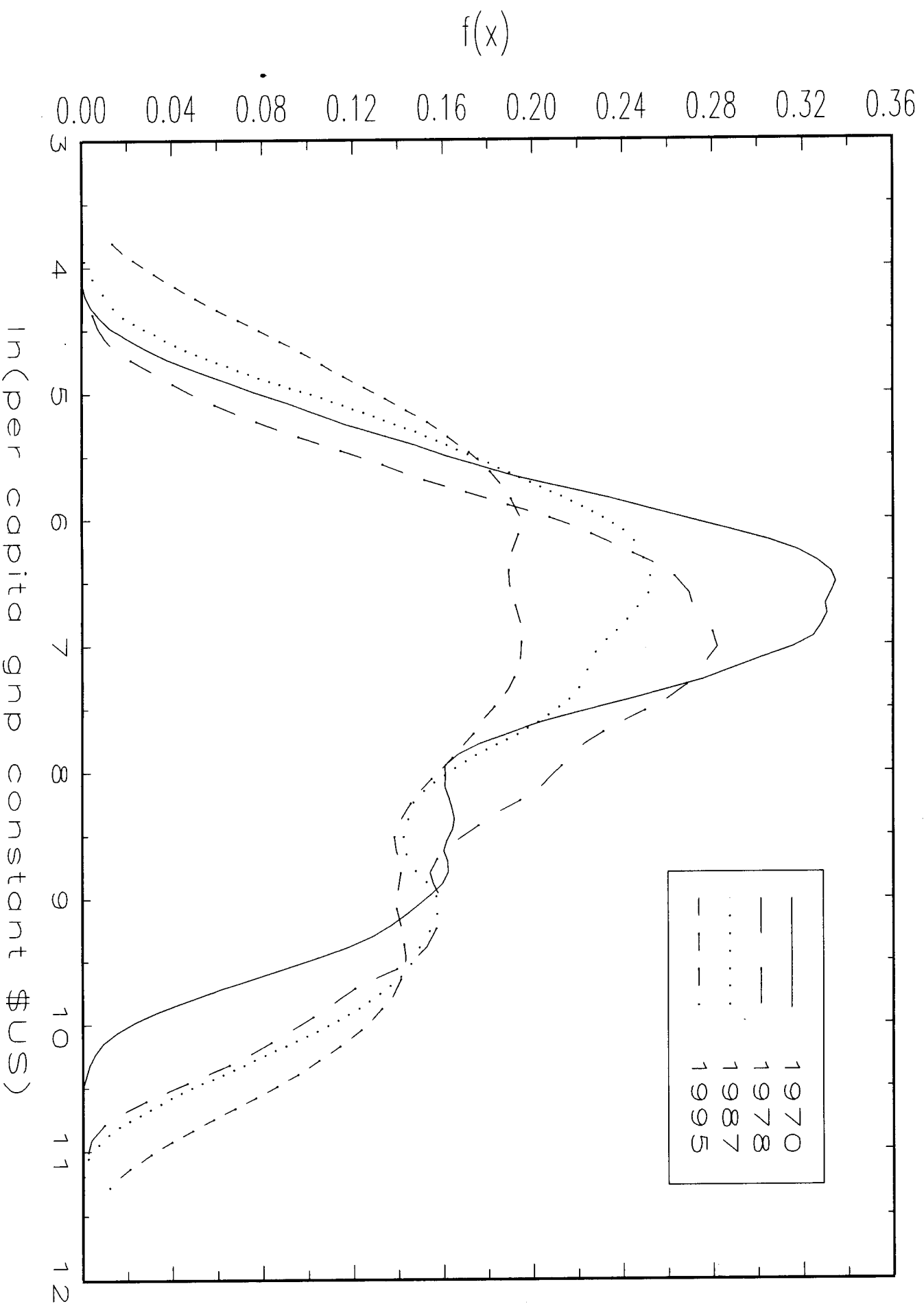
Population Weighted Per Capita GNP Distributions