A Polarization-Cohesion Perspective on Cross Country Convergence

Gordon Anderson*
University of Toronto

Oliver Linton†
The London School of Economics

Teng Wah Leo‡
St. Francis Xavier University

17 June, 2011

***Corresponding Author.** Address: Department of Economics, University of Toronto, 150 St. George Street, Toronto, Ontario, M5S 3G7, Canada. Email Address: anderson@chass.utoronto.ca

†Address: Department of Economics London School of Economics and Political Science, Houghton Street, London WC2A 2AE, United Kingdom. Email Address: o.linton@lse.ac.uk

‡Address: Department of Economics, St. Francis Xavier University, 1 West Street, Antigonish, Nova Scotia, B2G 2W5, Canada. Email Address: tleo@stfx.ca

Abstract

Understanding whether the gap between rich and poor country wellbeing is narrowing is really about whether rich and poor groups can be identified in the overall size distribution of the characteristic of interest, and how those respective subgroup size distributions are changing. Here two simple statistics for analyzing the issue are introduced which are capable of discerning, in many dimensions, changes in the underlying distributions which reflect combinations of increasing (decreasing) subgroup location differences and decreasing (increasing) subgroup spreads, which are the characteristics of polarization (de-polarization). When applied in an examination of the distribution of lifetime GDP per capita over time, the population weighted version exhibits de-polarization and the unweighted version exhibits polarization. As a collection of countries, Africa is diverging from the rest of the world regardless of the weighting scheme.

JEL Code: C14, I32, O47

Key Words: Polarization; Convergence; Overlap Measure; Trapezoid Measure

1 Introduction

There has been much debate over whether the gap between rich and poor countries' GDP per capita has narrowed, and the jury is still out as to whether differences between nations in this dimension has been reduced or not (Anand and Segal 2008). Often this debate has been pursued in terms of the nature of and change in inequality in the size distribution of country GDP per capita. Underlying this interest is a concern for wellbeing (and the lack of progress of the poor countries) which judges too much inequality as a bad thing. The argument has also transcended the use of simple GDP per capita measures extending the debate to broader measures that incorporate length and quality of life in the calculus[1], to an individualistic sense of income[2], and to more general indices of wellbeing such as the Human Development Index[3].

Following a growth regression literature which focused on Beta convergence (related to the coefficient on lagged income in a growth regression based upon variants of the Solow growth model) and Sigma convergence (related to the conditional variance of incomes), eventually culminating in Barro and Sala-i-Martin (1992), Mankiw, Romer and Weil (1992), and Galor (1996)[4], there has been extensive interest in examining the relative merits of the *Absolute Convergence* Hypothesis versus the *Club Convergence* Hypothesis[5]. The latter hypothesis corresponds to a tendency toward multiple modes (multiple equilibria) in the distribution of a country characteristic of interest (usually some measure of income per capita) and the former corresponds to a tendency to uni-modality (common equilibria) in that distribution with perhaps a diminution of its spread.

In other words, the issue is really about the changing nature of the anatomy of distributions of wellbeing indicators. With the Absolute Convergence Hypothesis concern is with the spread of or inequality in the distribution of interest, while with the Club Convergence Hypothesis the focus is the emergence of sub-group distributions (of which the overall distribution is a mixture) and the relative movements of those distributions. To

[1]Nordhaus (2002); Becker, Philipson and Soares (2005); Decancq, Decoster and Schokkaert (2009)

[2]Bourguignon and Morrisson (2002); Milanovic (2005); Sala-i-Martin (2006)

[3]Anderson (2005); Fleurbaey (2009); Fleurbaey and Gaulier (2009); Anderson, Crawford and Leicester (2010); Jones and Klenow (2010)

[4]All of which were comprehensively reviewed in Temple (1999)

[5]See for example Bianchi (1997), Jones (1997), Quah (1997), Paap and van Dijk (1998), Durlauf and Quah (1999), Johnson (2000), Islam (2003), Anderson (2004a), Beaudry, Collard and Green (2005), Durlauf and Fafchamps (2005), Pittau and Zelli (2006) and Durlauf, Johnson and Temple (2009)

illustrate how changes in the underlying distributions can affect the hypotheses, consider the limiting cases for how convergence between rich and poor groups could be brought about. Firstly, it can be brought about by diminishing within group identity (agents within groups becoming less alike) without any diminution of growth rate differentials between groups. Alternatively, it can be brought about by reductions in these differentials without any diminution of within group identity. This is the essence of the polarization literature initiated by Esteban and Ray (1994), Foster and Wolfson (1992), Wolfson (1994) and further developed in Anderson (2004a), Anderson (2004b) and Duclos, Esteban and Ray (2004), which distinguishes itself from pure notions of inequality since it can be readily shown that increased polarization can either reduce or increase inequality as conventionally measured. Here the convergence-divergence issue is examined empirically from a polarization-cohesion perspective

In addition, it may be argued that fundamental notions of individualistic welfare underlie much of the work in this area in the sense that it is the wellbeing of individuals in poor societies that are of concern with respect to their lack of economic growth relative to those in rich societies. In as much as this is the case, so that per capita aggregates represent the "average" agent in the economy, due consideration should be given to population weighting observations (for example the GDP per capita for China actually represents over 25% of the sample population, whereas that for Ireland represents less than 1%, making a strong argument for observations on those countries being viewed accordingly). On the other hand, if the life expectancy-GDP per capita nexus is viewed as a technological relationship (in the sense that the production function of individual wealth is a function of life expectancy and GDP per capita), and each country's realization is viewed as an observation on a particular technology blueprint so that interest is focused on the "average" technology, the argument for population weighting is much weaker.

These issues are addressed by employing new measures of convergence-divergence developed for the related literature on polarization. Rather than infer the nature of the ergodic distribution by using a regression technique that relates the elements of income distributions measured at two points in time (and assumes the relationship to be homogeneous across those elements), here the potential groups or clubs are (partially) identified via the anatomy of the distributions at two points in time and their relative progress is identified by measures of changes in that anatomy. Their attraction is that they have well understood statistical properties which avail us the opportunity of making inferences about the extent of convergence. The progress of GDP per capita and life expectancy of 123 countries over the period of 1990-2005 drawn from the World Bank's World Development Indicators (WDI) data set is considered, both with and without population weighting adjustments, and special consideration is given to the collection of African countries as a separate entity. One of the points to be made is that population weighting matters in that it makes a substantive difference to the results. It would be very easy to make the point by including China and India in the sample, since they have enjoyed growth rates well above the average over the sample period and constitute over a third of the population sample, and inevitably exacerbate the differences in weighted and unweighted results. For this reason they have been excluded from the analysis[6].

In the following, section 2 considers the links between the Convergence and Polarization literatures. Section 3 introduces the new measures and outlines their statistical properties. The application is reported in section 4 and conclusions are drawn in section 5.

2 Convergence and Polarization

Examining whether or not the poor versus rich nation divide is growing or diminishing within the context of *Polarization*, is really about eliciting from an observed mixture of distributions how the constituent sub-distributions (representing the respective "clubs") are behaving in terms of their movement relative to each other. In the Convergence literature, divergence has been associated with non-decreasing spread or inequality in the overall mixture distribution, while convergence associated with its non-increasing spread. The Polarization literature has been at pains to distinguish itself from pure inequality measurements. Following Esteban and Ray (1994), polarization between two groups is the consequence of a combination of two factors, increased within-group identification (usually associated with diminishing within-group variances, i.e. members of respective clubs coalesce), and increased between-group *alienation* (usually associated with increasing between-group differences in location, i.e. members of different clubs becoming more un-alike). Viewed in this context, global convergence or divergence is different from trends in the global variance, which is a monotonically increasing function of absolute between-

[6] In fact their inclusion does not alter the substantive results at all. The results are available from the authors upon request

group location differences and within-group dispersions, both of which can change in either direction with increased *polarization* (the Club Convergence Hypothesis).

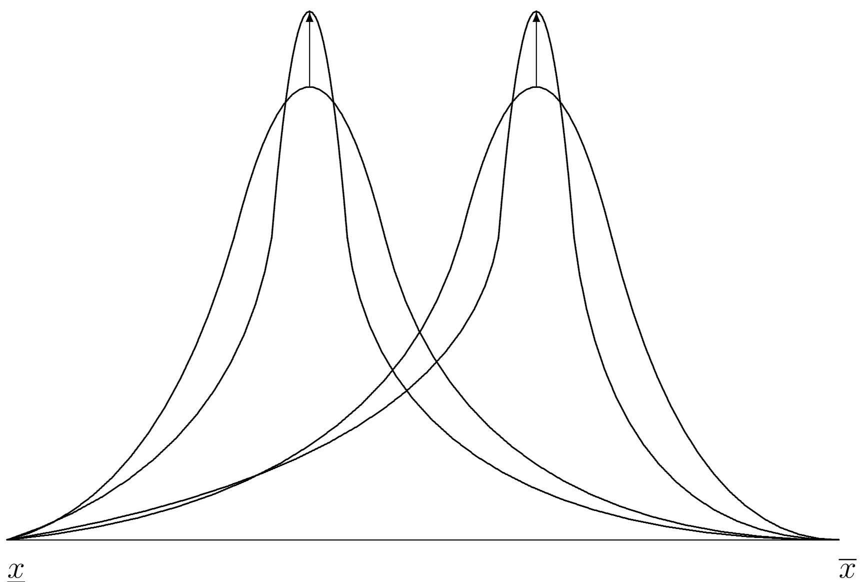
To see this, consider an equal weighted mixture of two normal distributions with equal variances, that is $x_1 \sim N(\mu_1, \sigma^2)$ and $x_2 \sim N(\mu_2, \sigma^2)$ are the sub-group or club distributions, and the mixture distribution becomes

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \frac{\left(\exp\left(\frac{(x-\mu_1)^2}{2\sigma^2}\right) + \exp\left(\frac{(x-\mu_2)^2}{2\sigma^2}\right)\right)}{2}$$

This distribution will be unimodal if $(\mu_1 - \mu_2)^2 < 27\sigma^2/8$ and will be bimodal (i.e. twin peaks will emerge) when $(\mu_1 - \mu_2)^2 > 27\sigma^2/8$, and has a variance of $((\mu_1 - \mu_2)^2 + 4\sigma^2)/4$ (see Johnson, Kotz and Balakrishnan (1994)). A move from a unimodal distribution to a bimodal distribution (consistent with the Club Convergence Hypothesis) that is the result of diminishing within sub-group variances (more homogeneous subgroup behavior) will be accompanied by diminishing variance in the population mixture (contrary to the Club Convergence hypothesis). Such a movement of the component sub-distributions is depicted in figure 1, where $x_1, x_2 \in [\underline{x}, \overline{x}]$. This is essentially polarization brought about by increasing within sub-group identity or association, rather than increased alienation between groups (see Duclos et al. (2004)), and is very much in the nature of within-group convergence and between-group divergence. Of course the reverse process will yield a trend toward unimodality with increasing variance (contrary to the Absolute Convergence Hypothesis). Furthermore, Anderson (2004a) showed that an alienation based polarization between two groups can be contrived, wherein location and spread preserving right skewing of the rich distribution and left skewing of the poor distribution will render polarization without any change in subgroup location and spread characteristics i.e. without any change in sub-group or global variance.

To reiterate, the examination of the Absolute versus Club Convergence Hypotheses thus boils down to examining the relative movements of the component sub-distributions. Observed changes in the variance of the overall mixture distribution can be a misleading statistic for verifying those hypotheses, since as has been demonstrated, convergence could engender movements in the variance of the mixture in either direction. Trends in the anatomy of the distribution of interest can be identified by polarization tests based on stochastic dominance relationships between the sub-distributions (Anderson 2004b), but it can be a cumbersome approach. This paper proposes two simple transparent statistics

Figure 1: Polarization with Decrease in Global Variance



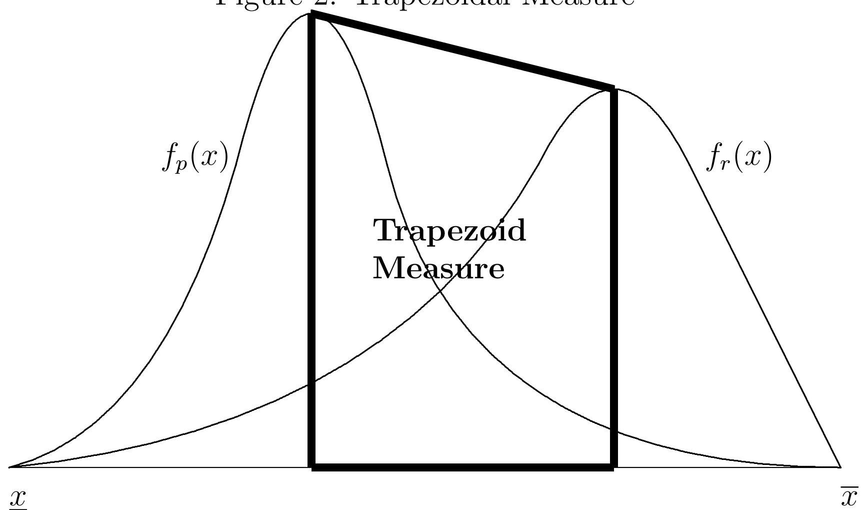
for inclusion into the researchers' toolkit, the Overlap measure and the Trapezoidal measure, changes in which reflect a combination of increasing (decreasing) sub-group location differences and decreasing (increasing) sub-group spreads, which are the characteristics of polarization (global convergence). The former statistic is only of use when the sub-distributions are known, while the latter can be used regardless of whether the sub-groups are known provided the distribution of interest is multi-modal.

3 The Methods

Suppose the rich and poor distributions are separately identified and let $x_{m,p}$ be the value of ln GDP per capita (x, where $x \in X$, and $X \subset \mathbb{R}$) at the modal point of the poor distribution $f_p(x)$, and $x_{m,r}$ the corresponding value for the rich distribution $f_r(x)$. In these circumstances, the area of the trapezoid formed by the heights of the distributions at their modal points and the distance between the two modal points provides a measure of the polarization or divergence of the poor and rich countries. The idea and measure is depicted in figure 2. Similarly the area of overlap of the distributions would also provide an index, provided there was an overlap (this is indeed a disadvantage of this technique since it is uninformative when the distributions are far apart or have minimal overlap), and is depicted in figure 3.

However, when the sub-distributions are not separately identified (by which is meant that the constituents of each group cannot be distinguished or separated from each other) but are embedded in a mixture, the Overlap measure is no longer useful, while fortunately the Trapezoid measure is, provided the mixture is bimodal (See figure 4 using data from the subsequent application). It is important to note that though these measures have been introduced in a univariate context, both are readily implemented when the distributions are multivariate in nature, a feature that will be exploited in the following application.

Figure 2: Trapezoidal Measure



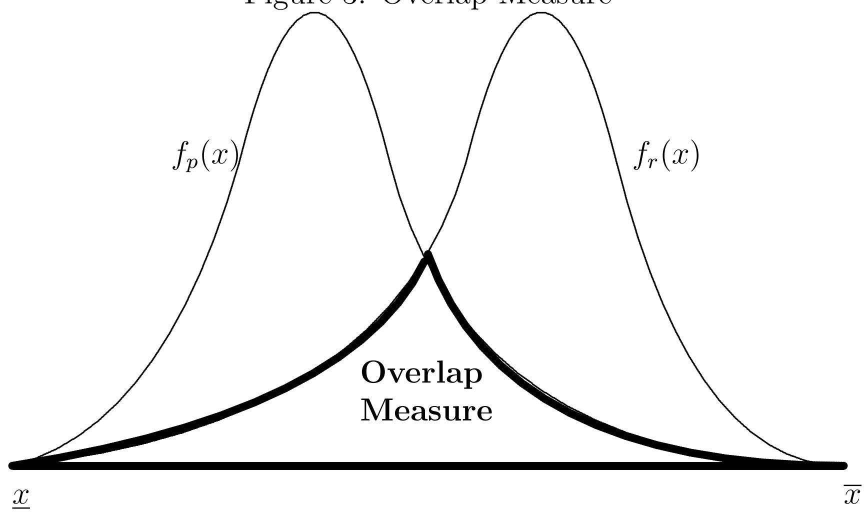
For two distributions $f_p(x)$ and $f_r(x)$, the Overlap measure (OV) is defined as:

$$OV = \int_{-\infty}^{\infty} \min\{f_p(x), f_r(x)\}dx \tag{1}$$

The distribution of this measure has been fully developed in Anderson, Linton and Whang (2009), where the contact set, its complements and corresponding probabilities are defined respectively as:

$$\begin{aligned} C_{f_p,f_r} &= \{x \in \mathbb{R} : f_p(x) = f_r(x) > 0\}; & p_0 &= \Pr(X \in C_{f_p,f_r}) \\ C_{f_p} &= \{x : f_p(x) < f_r(x)\}; & p_p &= \Pr(X \in C_{f_p}) \\ C_{f_r} &= \{x : f_p(x) > f_r(x)\}; & p_r &= \Pr(X \in C_{f_r}) \end{aligned}$$

Figure 3: Overlap Measure



The kernel estimator of the Overlap Index,

$$\widehat{OV} = \int_{-\infty}^{\infty} \min\left\{\widehat{f}_p(x), \widehat{f}_r(x)\right\} dx \tag{2}$$

where

$$\widehat{f}_p(x) = n^{-1}\sum_{i=1}^{n} K_b(x - X_i) \quad ; \quad \widehat{f}_r(x) = n^{-1}\sum_{i=1}^{n} K_b(x - Y_i)$$

$K_b(.) = K(./b)/b^d$, K is a s-times differentiable kernel function, b is the bandwidth sequence, d is the number of dimensions, and $s > d$. Anderson et al. (2009) showed (2) to be normally distributed of the form:

$$\sqrt{n}(\widehat{OV} - OV) - a_n \Longrightarrow N(0, v)$$

where

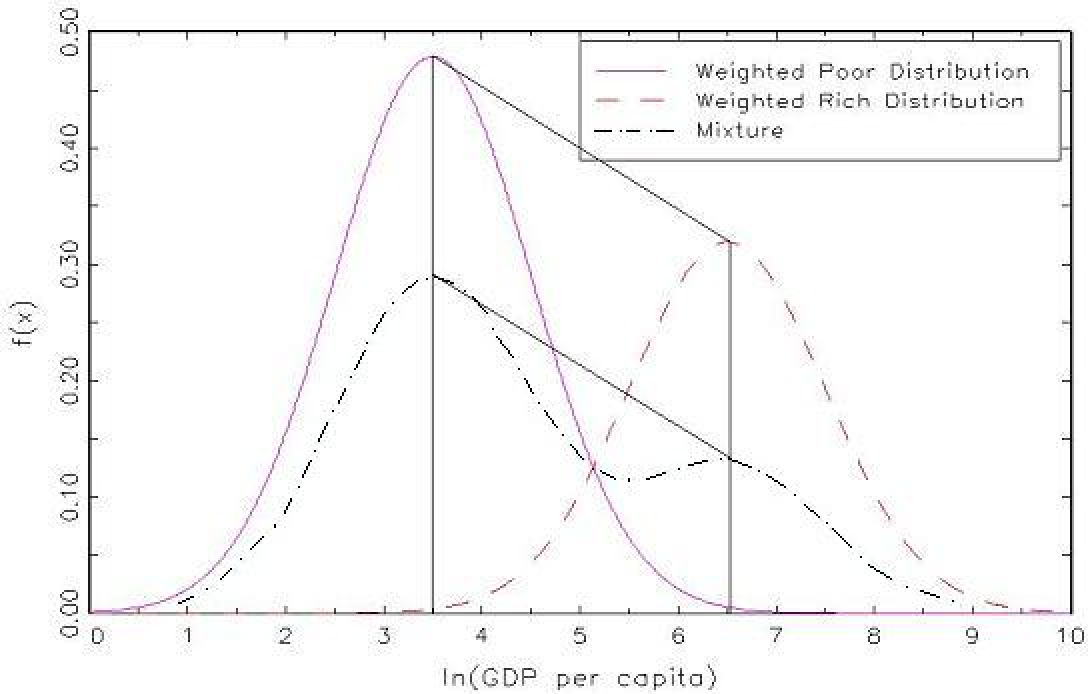
$$v = p_0\sigma_0^2 + p_p(1 - p_p) + p_r(1 - p_r)$$

a_n and σ_0^2 are bias correction factors (see Anderson et al. (2009) for details).

The intuition behind the Trapezoidal measure is borne out of the prior discussion on polarization[7]. Let $x_{m,i}$, $i = \{p, r\}$ be the modes of the respective poor and rich distributions, where $x_{m,i} \in X_i$, and $X_i \subset \mathbb{R}$. Firstly, since the intensity of within group association

[7]Furthermore, inspection of the Duclos et al. (2004) index reveals it to be proportionate to the expected value of the area of all trapezoids that can be formed under the density function, whose average height is $f(x)^\alpha$ and whose base is $|y - x|$.

Figure 4: 60-40 Weighted Sub-Distributions and the Mixture



is represented by the averaged heights of the modal points $f_p(x_{m,p})$ and $f_r(x_{m,r})$, following the intuition that the greater the mass within a region close to the modal point, the greater will the height of the density be. Secondly, since the modes themselves are measures of the centers of their respective groups around which constituent members are clustered, the use of the Euclidean distance between the two modal points representing the sense of alienation between two groups is only natural. It is interesting to speculate how the identity components could be interpreted. If I am poor, the poor modal height $f_p(x_{m,p})$ tells me the extent to which there are others like me or close to me, so that the higher it is the more identification with my group will I perceive. The rich modal height $f_r(x_{m,r})$ tells me how easily I can identify "the other club" and reflects how strongly I may perceive the other group from whom I am alienated. The higher the rich modal height the more closely associated the agents in that club are, the lower it is the more widely dispersed they are.

Formally, when the poor and rich distributions are separately identified in J dimensions, the trapezoidal measure $BIPOL$ may be written as:

$$BIPOL = \frac{1}{2}\left\{f_p(\mathbf{x}_{m,p}) + f_r(\mathbf{x}_{m,r})\right\}\frac{1}{\sqrt{J}}\sqrt{\sum_{j=1}^{J}\frac{(x_{m,p,j}-x_{m,r,j})^2}{\mu_j}} \tag{3}$$

where $\mathbf{x}_{m,i}$, $i = \{p, r\}$, is the vector of modes for group i with typical element $x_{m,i,j}$ on each dimension $j \in [1, J]$, and μ_j is the average of the modes from the two groups in the j'th dimension. When the groups are not known separately (denoted with subscript NI), the index is calculated from the modal points of the mixture distribution. To understand the variation, first note that the poor and rich modes may be written in terms of the underlying distributions as,

$$f(\mathbf{x}_{m,i}) \quad = \quad f_r(\mathbf{x}_{m,i}) + \omega\left(f_p(\mathbf{x}_{m,i}) - f_r(\mathbf{x}_{m,i})\right) \tag{4}$$

where $i = \{p, r\}$. Therefore, the index may be written as:

$$BIPOL_{NI} = \frac{1}{2}\left\{f(\mathbf{x}_{m,p}) + f(\mathbf{x}_{m,r})\right\}\left\{\frac{1}{\sqrt{J}}\sqrt{\sum_{j=1}^{J}\frac{(x_{m,p,j}-x_{m,r,j})^2}{\mu_j}}\right\} \tag{5}$$

The estimator of the trapezoid in the univariate case, when the underlying sub-distributions are not known separately, is given by:

$$\widehat{BIPOL} = \frac{1}{2}\left(\widehat{f}_p(\widehat{x}_{m,p}) + \widehat{f}_r(\widehat{x}_{m,r})\right)|(\widehat{x}_{m,p} - \widehat{x}_{m,r})| \tag{6}$$

Here $x_{m,i}$ is the mode for group i and $f_i(x_{m,i})$ is the value of the density at the modal point of group i, and hats refer to kernel estimators of the corresponding concepts. Appendix A.1 sketches the development of the distribution of $BIPOL$ in the univariate case as:

$$\begin{aligned} &(nh^3)^{1/2}(\widehat{BIPOL} - BIPOL) \\ \xrightarrow{D}\ \ & N\left(\text{Bias}, \tfrac{1}{4}\left\{f_r(x_{m,r}) + f_p(x_{m,p})\right\}^2\left\{\tfrac{f_r(x_{m,r})}{[f_r''(x_{m,r})]^2} + \tfrac{f_p(x_{m,p})}{[f_p''(x_{m,p})]^2}\right\}||K'||_2^2\right) \end{aligned} \tag{7}$$

The bias in $\widehat{BIPOL}$, the estimator of the trapezoid, is a consequence of the bias inherent in the estimate of $f_i(x_i)$, $i = \{p, r\}$. When x_i, $x_i \in X_i$, is a modal value, this bias can be expected to be of small order since, from Pagan and Ullah (1999) theorem 2.2, the bias in the kernel estimate of $f_i(x_i)$ at $x_i = x_{m,i}$ up to $O(h^2)$ is given by,

$$\frac{h^2}{2}\int \Psi^2 K(\Psi)d\Psi\frac{d^2 f_i(x_{m,i})}{dx_{m,i}^2}$$

where h is the window width, K is the kernel function and $\Psi = ((x_i - x_{m,i})/h)$. Since $x_{m,i}$ is a modal value of x_i, $\frac{d^2 f_i(x_{m,i})}{dx_{m,i}^2}$ is zero making the bias in $\widehat{BIPOL}$ of small order. Tests are based on the trapezoid measure being asymptotically normally distributed with a variance approximately equal to

$$\frac{1}{4}(f_r(x_{m,r}) + f_p(x_{m,p}))^2 \left(\frac{f_r(x_{m,r})}{[f_r''(x_{m,r})]^2} + \frac{f_p(x_{m,p})}{[f_p''(x_{m,p})]^2} \right) ||K'||_2^2$$

where K is the Gaussian kernel, and $||K'||_2^2$ is the L^2 norm of the first derivative of the Gaussian kernel function. Note that second derivatives of $f(.)$ can be estimated (again based upon a Gaussian kernel) as:

$$f^{(s)}(x) = \frac{(-1)^s}{nh^{s+1}} \sum_{i=1}^{n} K^{(s)} \left(\frac{x_i - x}{h} \right) \tag{8}$$

where for $s = 2$,

$$K^{(2)} \left(\frac{x_i - x}{h} \right) = \frac{\left[\left(\frac{x_i - x}{h} \right)^2 - 1 \right] e^{-0.5\left(\frac{x_i - x}{h} \right)^2}}{2\pi} \tag{9}$$

When the poor and rich distributions are not known separately, the principles remain the same. The primary problem that arises when it is not possible to separate the underlying distributions is whether the locations and heights of the two modes of the observed mixture distribution can be ascertained. Clearly the trapezoid test runs into trouble when the unidentified subgroups in the population are close to each other, which will be the case when the mixture is close to unimodal or the extent of the trough is somewhat limited. In this case the variability of the trapezoidal estimator will be inflated (since the terms $f''(.)$ in the variance formula will be close to 0), and thus changes in the trapezoid value will be hard to discern. Some discussion of modal detection is contained in the "Bump Hunting" literature reported in Silverman (1986), Bianchi (1997), and Henderson et al. (2008), but it is primarily in a univariate context. Among other approaches, extending the Dip test (Hartigan and Hartigan 1985) to multivariate contexts, alternative search methods (for example applying the Dip test along the predicted regression line) and parametric methods are all matters of current research. Nonetheless, for the application reported in the following section, the modes were easily ascertained.

4 The Application

Empirical growth models in the convergence literature have largely been concerned with poor country "catch-up" issues because of an underlying concern about the wellbeing (usually represented by the logarithm of GDP per capita) in those countries relative to rich countries. This particularly relates to the continent of Africa vis-à-vis the rest of the world, since Africa has the greatest proportion of "poor" countries. The illustrative application of the two statistics will likewise consider Africa (essentially Sub-Saharan Africa plus Morocco) and the Rest of the World as separate entities.

In terms of representing wellbeing, the use of ln(GDP per capita) involves two major issues. Firstly, growth regressions have very much a flavour of representative agent models with country i's ln(GDP per capita) being construed as the natural logarithm of consumption (or ln(income)) of the representative agent of the ith country. When used in unweighted growth regressions, the agent from Ireland (3.5 million population in 1990) has exactly the same weight as an individual in China (1,135 million in 1990), which is clearly inappropriate in the sense of an aggregate wellbeing measure. Secondly, microeconomic literature that built on Modigliani and Brumberg (1954) and Friedman (1957) developed models of agents who maximized the present value of lifetime wellbeing, $\left(\int_0^T U(C(t))e^{-\rho t}dt\right)$, subject to the present value of lifetime wealth, $\left(\int_0^T Y(t)e^{-rt}dt\right)$. Here, $U(.)$ is an instantaneous felicity function, Y is income, ρ is the representative agent's rate of time preference, and r is the market lending rate. Browning and Lusardi (1996) showed that this taken together with the assumption of constant relative risk aversion, and no bequest motive preference structure leads to a consumption smoothing model of the form:

$$C(t) = e^{\frac{(r-\rho)t}{\zeta}}C(0) \tag{10}$$

where ζ is the risk aversion coefficient, and by implication $g = (r-\rho)/\zeta$ is the consumption growth rate.

The latter point highlights the fact that the wellbeing of the representative agent depends upon her life expectancy, and since life expectancy varies considerably across countries, it needs to be accommodated in the calculus. If a country's GDP per capita at time "t" ($GDPpc(t)$) is thought to proxy average annual consumption over the lifetime of the representative agent (with life expectancy $T(t)$), and $\ln(GDPpc(t))$ is her instantaneous felicity function, with this happiness discounted to the present at the rate of time preference, ρ, then the wellbeing, $W(t)$, of the representative agent may be approximated as:

$$W(t) = \int_0^{T(t)} \ln(GDPpc(t))e^{-\rho s}ds = \frac{\ln(GDPpc(t))}{-\rho}(e^{-\rho T(t)} - 1) \tag{11}$$

Since ρ is not directly observable, for the current application, the rate used is the rate that each country's banking system offered on time deposits in 2005, a particularly low inflation year for most countries, and a year when most such rates were available in the World Bank data set[8]. Although this formulation of a wellbeing index is restrictive, it can nonetheless be generally agreed upon that wellbeing is some increasing bivariate function of GDP per capita and life expectancy, thereby justifying inferences regarding global convergence made using multivariate versions of the overlap and trapezoidal measures.

Table 1: Difference in Means Tests, GDP per capita & Life Expectancy, 1990 versus 2005

Population	GDP per capita		Life Expectancy		N
	T	$\Pr(t < T)$	T	$\Pr(t < T)$	
	Panel A: Unweighted				
All	-9.2122	0.0000	-3.8621	0.0000	123
Africa	-1.0587	0.1449	-1.486	0.0686	41
The Rest	-13.0865	0.0000	-18.1695	0.0000	82
	Panel B: Population Weighted				
All	-7.5185	0.0000	-7.5185	0.0000	123
Africa	-1.6787	0.0466	0.9692	0.8338	41
The Rest	-12.7969	0.0000	-12.7969	0.0000	82

The standard normal tests of the changes in GDP per capita and life expectancy for 1990 versus 2005 are presented in table 1, while the summary statistics of the data are reported in table A.1 in the appendix. From these tables it may be observed that there has been significant growth in GDP per capita and life expectancy over the observation period in the full sample, whether unweighted (Panel A) or population weighted (Panel B)[9]. However, the results for the African nations are not so clear cut, with all changes

[8] Actually the deposit rate was not available for 15 of the 123 countries, so rates in contiguous countries were averaged for the missing data.

[9] From table A.1 in the appendix, it may be observed that the variances have generally increased over the period between 1990-2005 (but not to a substantive extent) lending some support to the divergence

being insignificant at the 1% level, though it is interesting to note that sample weighting does affect the outcome of the life expectancy. Observe from table 1 that life expectancy fell in the unweighted sample, but rose in the weighted sample, while the results for the non-African countries reflect those of the full sample.

Table 2: Mixture Trapezoids

Year	Location/[Peak]		Trapezoid	Polarization Test $H_0 : BIPOL_{NI}^{1990} - BIPOL_{NI}^{2005} \geq 0$
	Unweighted			
1990	579.3366	107.6283	1.0880	
	[0.0009]	[0.0037]	(0.0185)	-1.0345
2005	616.0250	112.8695	1.1143	{0.1505}
	[0.0008]	[0.0036]	(0.0174)	
	Population Weighted			
1990	584.5778	144.3167	0.9738	
	[0.0011]	[0.0033]	(0.0174)	7.0454
2005	621.2662	201.9699	0.8152	{1.0000}
	[0.0009]	[0.0029]	(0.0143)	

Note: 1. Peak at the modes are in brackets and standard errors of the Trapezoid measure are in parenthesis. 2. t statistics are reported for the Polarization tests and $\Pr(t < T)$ values are in braces.

Kernel estimates of the univariate mixture distributions of lifetime wellbeing of equation (11), both unweighted and population weighted respectively, are reported in table 2, and depicted in figures 5 and 6[10]. It is immediately apparent that population weighting makes a considerable difference to the distribution's shape, emphasizing the bimodal nature of the distribution and suggesting a tendency for members of the poor group to have larger populations than members of the rich group.

The trapezoidal tests reported in table 2 highlight the impact of population weighting, with the unweighted sample only weakly supporting the null of convergence (i.e. it fails to reject the null), while the weighted sample strongly supports the hypothesis of convergence. This no doubt reflects Sala-i-Martin's (2006) finding of convergence when global

hypothesis.

[10]It is of interest to see how different are the pure GDP per capita distributions which are depicted in the appendix.

inequality is treated in an individualistic sense (population weighted) as opposed to when it is addressed in a between country (population unweighted) sense. The important point to stress here is that it is population weighting that has made the profound difference, and neither China nor India was included (which would have emphasized the difference).

Figure 5: Unweighted Mixture Distribution, 1990-2005

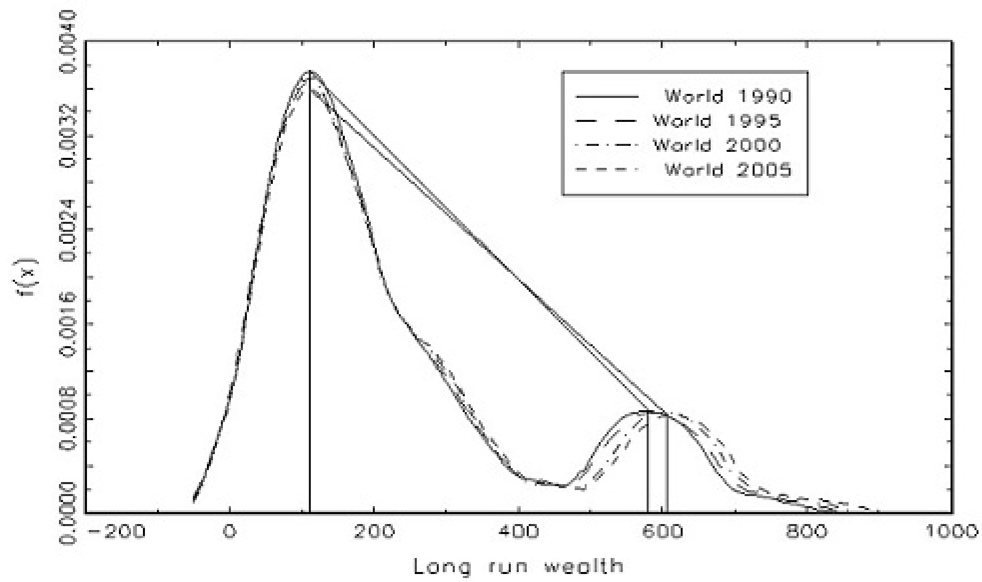


Figure 6: Population Weighted Mixture Distributions, 1990-2005

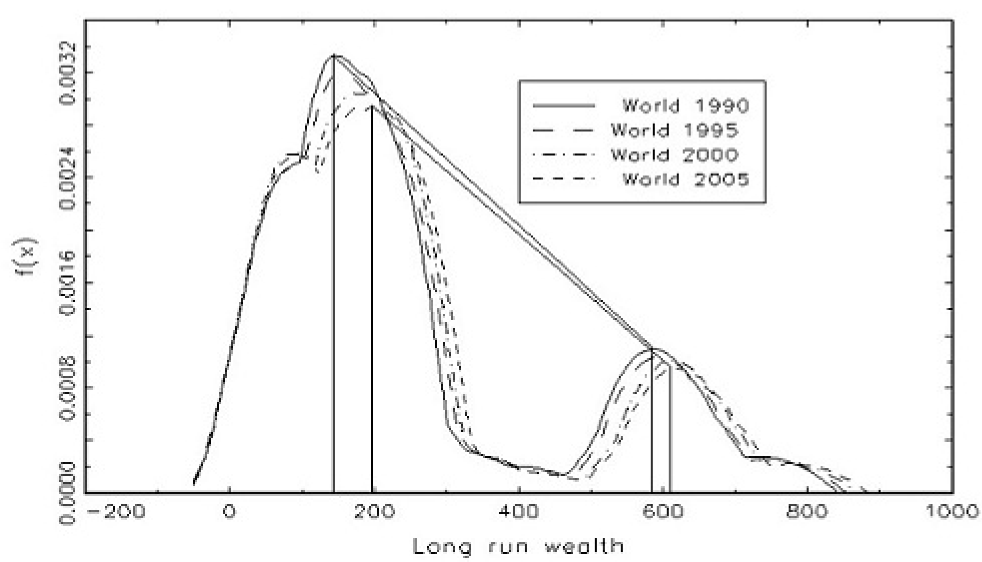


Table 3: Trapezoid Measure, Africa and the Rest Comparisons

	Panel A				Panel B			
	Africa vs. Other Poor Nations, Unweighted				Africa vs. Other Poor Nations, Weighted			
	Location/[Peak]		Trapezoid	Polarization Test $H_0: BIPOL_{NI}^{1990} - BIPOL_{NI}^{2005} \geq 0$	Location/[Peak]		Trapezoid	Polarization Test $H_0: BIPOL_{NI}^{1990} - BIPOL_{NI}^{2005} \geq 0$
1990	123.4107	165.3367	0.2204		91.9661	160.0960	0.3799	
	[0.0080]	[0.0025]	(0.0467)	-1.2651	[0.0084]	[0.0028]	(0.0589)	-1.0559
2005	118.1699	175.8183	0.3042	{0.1029}	91.9661	175.8183	0.4678	{0.1455}
	[0.0082]	[0.0024]	(0.0469)		[0.0084]	[0.0027]	(0.0589)	
	Africa vs. Other Non-Poor Nations, Unweighted				Africa vs. Other Non-Poor Nations, Weighted			
	Location/[Peak]		Trapezoid	Polarization Test $H_0: BIPOL_{NI}^{1990} - BIPOL_{NI}^{2005} \geq 0$	Location/[Peak]		Trapezoid	Polarization Test $H_0: BIPOL_{NI}^{1990} - BIPOL_{NI}^{2005} \geq 0$
1990	123.4107	574.1159	2.0281		91.9661	610.8012	2.4096	
	[0.0080]	[0.0010]	(0.0370)	-3.0467	[0.0084]	[0.0009]	(0.0448)	-2.2098
2005	118.1699	600.3197	2.1884	{0.0012}	91.9661	642.2458	2.5492	{0.0136}
	[0.0082]	[0.0009]	(0.0375)		[0.0084]	[0.0008]	(0.0446)	
	Africa vs. Other Nations				Africa vs Other Nations			
Year			Overlap Index (OV)	Difference in OV Test $H_0: OV^{1990} - OV^{2005} \leq 0$			Overlap Index (OV)	Difference in OV Test $H_0: OV^{1990} - OV^{2005} \leq 0$
1990			0.4551				0.4753	
			(0.0953)	0.2151			(0.0955)	0.3043
2005			0.4262	{0.5852}			0.4343	{0.6196}
			(0.0946)				(0.0948)	

Note: 1. Mass/density at the modes are in brackets and standard errors of the Trapezoid and Overlap measures are in parenthesis.

2. t statistics are reported for the Polarization and Difference in OV tests, and $\Pr(t < T)$ values are in braces for both tests.

Figure 7: Unweighted Distribution for Africa and the Rest, 1990-2005

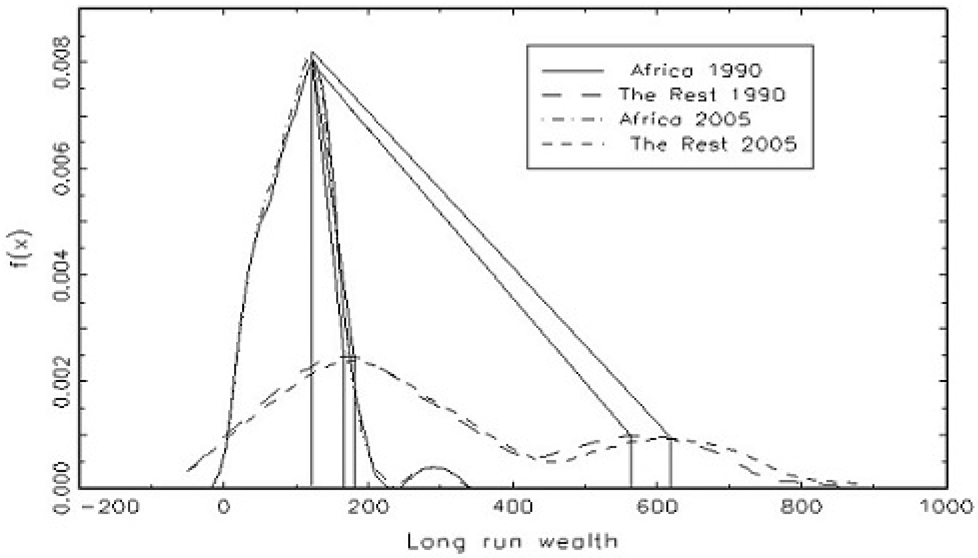
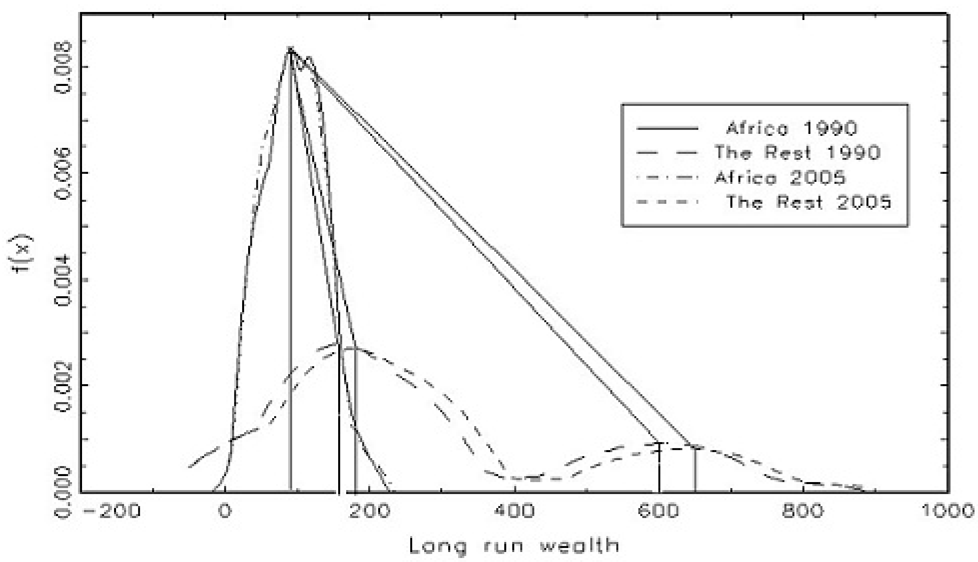


Figure 8: Population Weighted Distribution for Africa and the Rest, 1990-2005



Turning to a comparison of Africa and the Rest of the World, figures 7 and 8 depict the kernel estimates of the unweighted and population weighted distributions respectively. The results of the trapezoidal and overlap measures are reported in table 3. Here convergence between Africa and other Poor Countries in the world, reported in panel A, cannot be rejected, whereas convergence between Africa and other Non-Poor Countries can be rejected in both unweighted and weighted distributions as observed in panel B. This is not supported by the Overlap measure of panel C, which fails to reject the null hypothesis of convergence between Africa and the rest of the world as a whole.

Table 4: Within Distributions Polarization Trapezoids for the Joint Distribution of log GDP per Capita & Life Expectancy, All Countries

	Panel A: Unweighted			
	Location/[Peak]		Trapezoid	$\left(H_0 : BIPOL_{NI}^{1990} - BIPOL_{NI}^{2005} \geq 0\right)$
1990	6.5984,57.8648	9.0104,72.1639	0.0685	
	[0.0049]	[0.0046]	(0.0042)	1.8974
2005	6.7338,59.9339	9.0205,75.009	0.0580	{0.9711}
	[0.0037]	[0.0039]	(0.0036)	
	Panel B: Population Weighted			
	Location/[Peak]		Trapezoid	$\left(H_0 : BIPOL^{1990} - BIPOL^{2005} \geq 0\right)$
1990	6.0936,59.7713	9.0104,72.1639	0.0432	
	[0.0068]	[0.0042]	(0.0024)	3.3598
2005	6.6702,65.5871	9.0205,74.5379	0.0288	{0.9996}
	[0.0062]	[0.0041]	(0.0036)	

Note: 1. The peak of the mode are in brackets and standard errors of the Trapezoid are in parenthesis. 2. t statistics are reported for the Polarization and Difference in OV tests, and $\Pr(t < T)$ values are in braces for both tests.

As suggested above, if the very restrictive version of the univariate welfare measure is of concern, the robustness of the univariate result for global convergence in income and life expectancy can be addressed in a multivariate framework by treating GDP per capita, and life expectancy as separate variables as opposed to the lifetime wellbeing measure of equation 11, using both the Trapezoidal and Overlap measures. This approach has the additional benefit of not imposing a functional relationship on the characteristics of interest. The results for within and between distribution trapezoids are reported in tables 4 and 5 respectively. The within distribution results in table 4 reveal that the null of convergence cannot be rejected at all usual levels of significance. However, in table 5 the null of convergence is unanimously rejected for both the Trapezoid and Overlap measures. On the whole, this suggests that evidence of global convergence is driven by improvements in wellbeing amongst other poorer nations outside of Africa.

Table 5: Between Distribution Polarization Trapezoids & Overlap Measure for the Joint Distribution of log GDP per Capita & Life Expectancy, Africa and the Rest

	Unweighted Distributions		Weighted Distributions	
	1990	2005	1990	2005
Trapezoid Measure	1.5284	1.7921	0.3653	0.6925
	(0.1225)	(0.0992)	(0.0683)	(0.0975)
t Statistic for Difference	-1.6731		-2.7480	
$\left(H_0 : BIPOL^{1990} - BIPOL^{2005} \geq 0\right)$	{0.0472}		{0.0030}	
Overlap Measure	0.2516	0.1298	0.5707	0.3839
	(0.0488)	(0.0241)	(0.0407)	(0.0241)
t Statistic for Difference	2.2365		3.9517	
$\left(H_0 : OV^{1990} - OV^{2005} \leq 0\right)$	{0.9873}		{1.0000}	

Note: 1. The peak of the modes are in brackets and standard errors of the Trapezoid and Overlap measures are in parenthesis. 2. t statistics are reported for the Polarization and Difference in OV tests, and $\Pr(t < T)$ values are in braces for both tests.

5 Conclusion

Cross-country examinations of global convergence are about the changing nature of the anatomy of distributions of wellbeing indicators. As limiting cases, separately identified rich and poor group convergence can be brought about by diminishing within-group identity (agents within groups becoming less alike) without any diminution of group growth rate differentials (group locations converging) or it can be brought about by group locations converging (diminishing between-group alienation) without any diminution of within-group identity. Further, the diminution of within-group identity increases global variance, whereas diminishing between-group alienation reduces global variance. Given this myriad of possibilities regarding the behavior of distributions, it is argued here that measures of the extent to which distributions of wellbeing indicators overlap (the Overlap measure), or measures which are monotonically increasing functions of the extent of a distribution's modality, and the extent to which their modal coordinates differ (the Trapezoid measure), provide reliable instruments for identifying trends in global convergence.

Two additional issues need to be addressed as well. Firstly, in the global convergence calculus, the choice of whether the concern is convergence in individualistic wellbeing or convergence in wellbeing producing technologies is important. If the former, an inter-country analysis requires consideration of population weighting issues, while the latter does not. The second issue that needs to be addressed is to consider convergence in terms of lifetime wellbeing based upon some combination of measures of annual expected income and life expectancy. This concern is however addressed by both measures since they are amenable to multivariate analysis. Further the measures' nonparametric nature negates concerns regarding the appropriate functional relationships between the variables.

The issue of the changing nature of the rich country-poor country divide has been addressed here by introducing measures which reflect these considerations, and which have well-defined statistical properties. The attraction of this is that statistical inferences can be made as to the "significance" or not of the nature of convergence, whether it be in a multivariate or univariate paradigm. The indicators appear to work well in both single variable and multiple variable environments.

The results of the application indicate that including life expectancy in the calculus changes the results substantially, exacerbating our impression of Africa's relative plight and changing the shape of the distribution of wellbeing however measured. Likewise population weighting also changes the results substantially. While there appears to be poor country-rich country convergence in the world wellbeing distribution when considered on an individualistic basis, there is evidence of divergence when country data are viewed as observations on technologies. When Africa is separated out, it seems to be diverging from the rest of the world whether measured in an individualistic or technological sense.

References

Anand, Sudhir and Paul Segal, "What do we know about Global Income Inequality?," *Journal of Economic Literature*, 2008, *46* (1), 57–94.

Anderson, Gordon, "Making Inferences about the Welfare Poverty and Polarization of Nations: A Study of 101 Countries 1970-1995," *Journal of Applied Econometrics*, 2004, *19*, 537–550.

____ , "Towards an Empirical Analysis of Polarization," *Journal of Econometrics*, 2004, *122*, 1–26.

____ , "Life Expectancy, Welfare and Inequality: The Example of Africa in the 1990s," *Review of Income and Wealth*, 2005, *51* (3), 455–468.

____ **, Ian Crawford, and Andrew Leicester**, "Welfare Rankings from Multivariate Data, a Nonparametric Approach," *Journal of Public Economics*, 2010. Forthcoming. Available online 9 September 2010.

____ **, Oliver Linton, and Yoon-Jae Whang**, "Nonparametric Estimation of a Polarization Measure," 2009. Mimeo Department of Economics, University of Toronto.

Barro, Robert J. and Xavier Sala-i-Martin, "Convergence," *Journal of Political Economy*, 1992, *100* (2), 223–251.

Beaudry, Paul, Fabrice Collard, and David A. Green, "Changes in the World Distribution of Output per Capita 1960-1998: How a Standard Decomposition Tells an Unorthodox Story," *Review of Economics and Statistics*, 2005, *87* (4), 741–753.

Becker, Gary S., Tomas J. Philipson, and Rodrigo R. Soares, "The Quantity and Quality of Life and the Evolution of World Inequality," *American Economic Review*, 2005, *95* (1), 277–291.

Bianchi, Marco, "Testing for Convergence: Evidence from Non-Parametric Multimodality Tests," *Journal of Applied Econometrics*, 1997, *12* (4), 393–409.

Bourguignon, Francois and Christian Morrisson, "Inequality Among World Citizens: 18201992," *American Economic Review*, 2002, *92* (4), 727–744.

Browning, Martin and Annamaria Lusardi, "Household Saving: Micro Theories and Micro Facts," *Journal of Economic Literature*, 1996, *34* (4), 1797–1855.

Decancq, Koen, Andre Decoster, and Erik Schokkaert, "The Evolution of World Inequality in Well-being," *World Development*, 2009, *37* (1), 11–25.

Duclos, Jean-Yves, Joan-Maria Esteban, and Debraj Ray, "Polarization: Concepts, Measurement, Estimation," *Econometrica*, 2004, *72*, 1737–1772.

Durlauf, Steven N. and Danny T. Quah, "The new empirics of economic growth," in J. B. Taylor and M. Woodford, eds., *Handbook of Macroeconomics*, Vol. 1 of *Handbook of Macroeconomics*, Elsevier, 1999, chapter 4, pp. 235–308.

—— **and Marcel Fafchamps**, "Social Capital," in Philippe Aghion and Steven Durlauf, eds., *Handbook of Economic Growth*, Vol. 1 of *Handbook of Economic Growth*, Elsevier, 2005, chapter 26, pp. 1639–1699.

—— **, Paul A. Johnson, and Jonathan R.W. Temple**, "The Econometrics of Convergence," in Terence C. Mills and Kerry Patterson, eds., *Palgrave Handbook of Econometrics*, Vol. 2 of *Applied Econometrics*, Elsevier, 2009.

Eddy, William F., "Optimum Kernel Estimators of the Mode," *The Annals of Statistics*, 1980, *8* (4), 870–882.

Esteban, Joan-Maria and Debraj Ray, "On the Measurement of Polarization," *Econometrica*, 1994, *62* (4), 819–851.

Fleurbaey, Marc, "Beyond GDP: The Quest for a Measure of Social Welfare," *Journal of Economic Literature*, 2009, *47* (4), 1029–1075.

—— **and Guillaume Gaulier**, "International Comparisons of Living Standards by Equivalent Incomes," *Scandinavian Journal of Economics*, 2009, *111* (3), 597–624.

Foster, James E. and Michael C. Wolfson, "Polarization and the Decline of the Middle Class: Canada and the U.S.," 1992. Mimeo, Vanderbilt University.

Friedman, Milton, *A Theory of the Consumption Function*, Princeton, New Jersey: Princeton University Press, 1957.

Galor, Oded, "Convergence? Inferences from Theoretical Models," *The Economic Journal*, 1996, *106* (437), 1056–1069.

Hartigan, J. A. and P. M. Hartigan, "The Dip Test of Unimodality," *Annals of Statistics*, 1985, *13*, 70–84.

Henderson, Daniel J., Christopher F. Parmeter, and Robert Russell, "Modes, Weighted Modes, and Calibrated Modes: Evidence of Clustering Using Modality Tests," *Journal of Applied Econometrics*, 2008, *23* (5), 607–638.

Islam, Nazrul, "What have We Learnt from the Convergence Debate?," *Journal of Economic Surveys*, 2003, *17* (3), 309–362.

Johnson, Norman L., Samuel Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, 2 ed., Vol. 1 of *Wiley Series in Probability and Mathematical Statistics* 1994.

Johnson, Paul A., "A Non-Parametric Analysis of Income Convergence Across the US States," *Economics Letters*, 2000, *69* (2), 219–223.

Jones, Charles I., "On the Evolution of the World Income Distribution," *The Journal of Economic Perspectives*, 1997, *11* (3), 19–36.

___ **and Peter J. Klenow**, "Beyond GDP? Welfare Across Countries and Time," *NBER Working Paper 16352*, 2010.

Mankiw, Gregory, David Romer, and David N. Weil, "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics*, 1992, *107* (2), 407–437.

Mann, H.B. and A. Wald, "On Stochastic Limit and Order Relationships," *The Annals of Mathematical Statistics*, 1943, *14* (3), 217–226.

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

Modigliani, Franco and Richard Brumberg, "Utility Analysis and the Consumption Function: An Interpretation of Cross-Section Data," in Kenneth K. Kurihara, ed., *Post-Keynesian Economics*, Rutgers University Press, 1954.

Nordhaus, William D., "The Health of Nations: The Contribution of Improved Health to Living Standards," *NBER Working Paper No. 8818*, 2002.

Paap, Richard and Herman K. van Dijk, "Distribution and Mobility of Wealth of Nations," *European Economic Review*, 1998, *42* (7), 135–165.

Pagan, Adrian and Aman Ullah, *Nonparametric Econometrics*, Cambridge: Cambridge University Press, 1999.

Pittau, Maria Grazia and Roberto Zelli, "Empirical Evidence of Income Dynamics across EU Regions," *Journal of Applied Econometrics*, 2006, *21* (5), 605–628.

Quah, Danny T., "Empirics for Growth and Distribution: Stratification, Polarization, and Convergence Clubs," *Journal of Economic Growth*, 1997, *2* (1), 27–59.

Sala-i-Martin, Xavier, "The World Distribution of Income: Falling Poverty and ... Convergence, Period," *Quarterly Journal of Economics*, 2006, *121* (2), 351–397.

Silverman, B. W., *Density Estimation for Statistics and Data Analysis*, New York: Chapman and Hall, 1986.

Temple, Jonathan, "The New Growth Evidence," *Journal of Economic Literature*, 1999, *37* (1), 112–156.

Wolfson, Michael C., "When Inequalities Diverge," *American Economic Review Papers and Proceedings*, 1994, *84* (2), 353–358.

A Appendix

A.1 Asymptotic Distribution of the Trapezoid Estimator

Many variants of this index are possible. Note the weights given to either the within group association or the between group alienation components could be varied if such emphasis is desired. Thus a general form of $BIPOL$ could be $(\text{Height}^{\alpha}\text{Base}^{1-\alpha})^2$, where $0 < \alpha < 1$ represents the relative importance of the self identification component. Similarly the modal point height components could be individually re-weighted to reflect the different importance of the identification component of the rich and poor groups. Note also that if indices based upon different numbers of characteristics are being compared, the identification component of the index should be scaled by the number of characteristics being contemplated based upon the fact that the peak of the joint density of J independent $N(0,1)$ is $1/\sqrt{J}$ times the height of one $N(0,1)$.

The estimator of the trapezoid is,

$$\widehat{BIPOL} = \frac{1}{2}\left(\widehat{f}_p(\widehat{\mathbf{x}}_{m,p}) + \widehat{f}_r(\widehat{\mathbf{x}}_{m,r})\right)\frac{1}{\sqrt{J}}\sqrt{\sum_{j=1}^{J}\frac{(\widehat{x}_{m,p,j} - \widehat{x}_{m,r,j})^2}{\widehat{\mu}_j}} \tag{A-1}$$

where μ_j is the average of the modes in the j'th dimension and where $\mathbf{x}_{m,i}$ is the modal vector for the i'th group, $i \in \{p, r\}$, with typical elements $x_{m,i,j}$ $j \in \{1, 2, ..., J\}$. Let hats denote the empirical counterparts to the population densities and values, so that $\widehat{f}_i$, $i \in \{p, r\}$, refer to the kernel estimates of the population density function.

Considering first the one-dimensional version,

$$\widehat{BIPOL} = \frac{1}{2}\left(\widehat{f}_p(\widehat{x}_{m,p}) + \widehat{f}_r(\widehat{x}_{m,r})\right)|(\widehat{x}_{m,p} - \widehat{x}_{m,r})| \tag{A-2}$$

Let K be a real valued Kernel function, h be the bandwidth, and n is the number of observations in the sample. We know from corollary 2.2 of Eddy (1980),

$$(nh^3)^{\frac{1}{2}}(\widehat{x}_{m,i} - x_{m,i}) \xrightarrow{D} N\left(\text{Bias}, \frac{f_i(x_{m,i})}{[f_i''(x_{m,i})]^2}||K'||_2^2\right) \tag{A-3}$$

where $||K'||_2^2$ is the L^2 norm of the first derivative of the Kernel function. Next, we write

$$\widehat{f}_i(\widehat{x}_{m,i}) - f_i(x_{m,i}) = \left(\widehat{f}_i(\widehat{x}_{m,i}) - \widehat{f}_i(x_{m,i})\right) + \left(\widehat{f}_i(x_{m,i}) - f_i(x_{m,i})\right) \tag{A-4}$$

Focusing on the first term on the right-hand side, let $b = (nh^3)^{-\frac{1}{2}}$ and define the random process $Z_n(t)$ as,

$$Z_n(t) = b^{-2}\left[\widehat{f}_i(x_{m,i} + bt) - \widehat{f}_i(x_{m,i})\right] \tag{A-5}$$

where $t \in [-T, T]$ for $T < \infty$. By Theorem 2.1 in Eddy (1980)

$$Z_n(t) \Longrightarrow Z(t) = \frac{f''(x_{m,i})}{2}t^2 + (-1)^{q+1}\frac{f^{(q+1)}(x_{m,i})}{q!}dB_q t + Yt$$

where Y is a normally distributed random variable, $N(0, f(x_{m,i})||K'||_2^2)$, $q \geq 2$ is an integer, $\lim_{n\to\infty}(nh^{3+2q})^{\frac{1}{2}} = d$, $f_i^{(q+1)}$ is the $q+1$'th order derivative and B_q is just the q'th moment of the kernel function. Then by the continuous mapping theorem (Mann and Wald 1943), it follows that,

$$b^{-2}\left[\widehat{f}_i(\widehat{x}_{m,i}) - \widehat{f}_i(x_{m,i})\right] = Z_n(\widehat{t}) \Longrightarrow Z(\widetilde{t}) \tag{A-6}$$

where $\widehat{t} = (nh^3)^{\frac{1}{2}}(\widehat{x}_{m,i} - x_{m,i})$ and $\widetilde{t} \sim N\left(\text{Bias}_t, \frac{f_i(x_{m,i})}{[f_i''(x_{m,i})]^2}||K'||_2^2\right)$. The bias term is,

$$(-1)^q\frac{d}{q!}\frac{f_i^{(q+1)}(x_{m,i})}{f_i''(x_{m,i})}B_q$$

Therefore,

$$\left[\widehat{f}_i(\widehat{x}_{m,i}) - \widehat{f}_i(x_{m,i})\right] = O_p(n^{-1}h^{-3}) \tag{A-7}$$

For the second term, note that by Theorem 2.6 in Pagan and Ullah (1999), pointwise at $x_{m,i}$,

$$\widehat{f}_i(x_{m,i}) - f_i(x_{m,i}) = o_p(1) \tag{A-8}$$

So that equation (A-4) is,

$$\widehat{f}_i(\widehat{x}_{m,i}) - f_i(x_{m,i}) = O_p(n^{-1}h^{-3}) \tag{A-9}$$

and it is non-normal. However, when $x_{m,i} \neq x_{m,j}$, $i \neq j, i, j \in \{p, r\}$, we have,

$$(nh^3)^{\frac{1}{2}}(|\widehat{x}_{m,i} - \widehat{x}_{m,j}| - |x_{m,i} - x_{m,j}|) \xrightarrow{D} N\left(\text{Bias}, \left\{\frac{f_i(x_{m,i})}{[f_i''(x_{m,i})]^2} + \frac{f_i(x_{m,j})}{[f_i''(x_{m,j})]^2}\right\}||K'||_2^2\right) \tag{A-10}$$

which is the dominant term in the limiting distribution. Note that the bias term is,

$$\sum_{i\in\{r,p\}}(-1)^q\frac{d}{q!}\frac{f_i^{(q+1)}(x_{m,i})}{f_i''(x_{m,i})}B_q$$

It follows that in this regular case,

$$
\begin{aligned}
& (nh^3)^{1/2}(\widehat{BIPOL} - BIPOL) \\
\xrightarrow{D} \; & N\left(\text{Bias}, \frac{1}{4}\left\{f_i(x_{m,i}) + f_j(x_{m,j})\right\}^2 \left\{\frac{f_i(x_{m,i})}{[f_i''(x_{m,i})]^2} + \frac{f_j(x_{m,j})}{[f_j''(x_{m,j})]^2}\right\} ||K'||_2^2\right)
\end{aligned}
\tag{A-11}
$$

On the other hand, when $x_{m,i} = x_{m,j}$, we will have half normal asymptotics.

A.2 Countries in the sample

Angola, Argentina, Armenia, Australia, Austria, Bahrain, Bangladesh, Belgium, Belize, Benin, Bhutan, Bolivia, Botswana, Brazil, Burkina Faso, Burundi, Cameroon, Canada, Cape Verde, Central African Republic, Chad, Chile, Colombia, Comoros, Congo, Dem. Rep.,Congo, Rep., Costa Rica, Cote d'Ivoire, Denmark, Djibouti, Dominican Republic, Ecuador, El Salvador, Ethiopia, Finland, France, Gabon, Gambia, Germany, Ghana, Greece, Guatemala, Guinea, Guinea-Bissau, Guyana, Haiti, Honduras, Hong Kong (China), Iceland, India, Indonesia, Iran, Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kenya, Korea, Lao PDR, Lebanon, Lesotho, Liberia, Lithuania, Luxembourg, Madagascar, Malawi, Malaysia, Mali, Malta, Mauritania, Mauritius, Mexico, Micronesia (Fed. Sts.), Morocco, Namibia, Nepal, Netherlands, New Zealand, Nicaragua, Niger, Nigeria, Norway, Pakistan, Panama, Paraguay, Peru, Philippines, Poland, Portugal, Rwanda, Samoa, Sao Tome and Principe, Saudi Arabia, Senegal, Sierra Leone, Singapore, Solomon Islands, South Africa, Spain, Sri Lanka, St. Vincent and the Grenadines, Sudan, Suriname, Swaziland, Sweden, Switzerland, Syria, Tanzania, Thailand, Togo, Tonga, Trinidad and Tobago, Uganda, United Arab Emirates, United Kingdom, United States, Uruguay, Vanuatu, Venezuela, Vietnam, Yemen, Zambia.

A.3 Pure GDP Per Capita Distributions

Figure A.1: Africa and the Rest, 1990

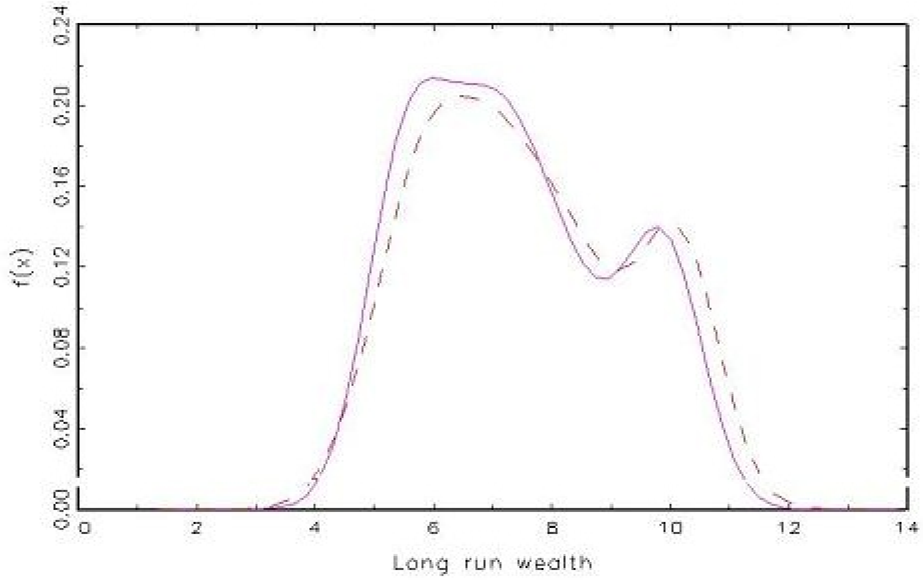


Figure A.2: Population Weighted Distribution of per Capita GDP for the World, 1990-2005

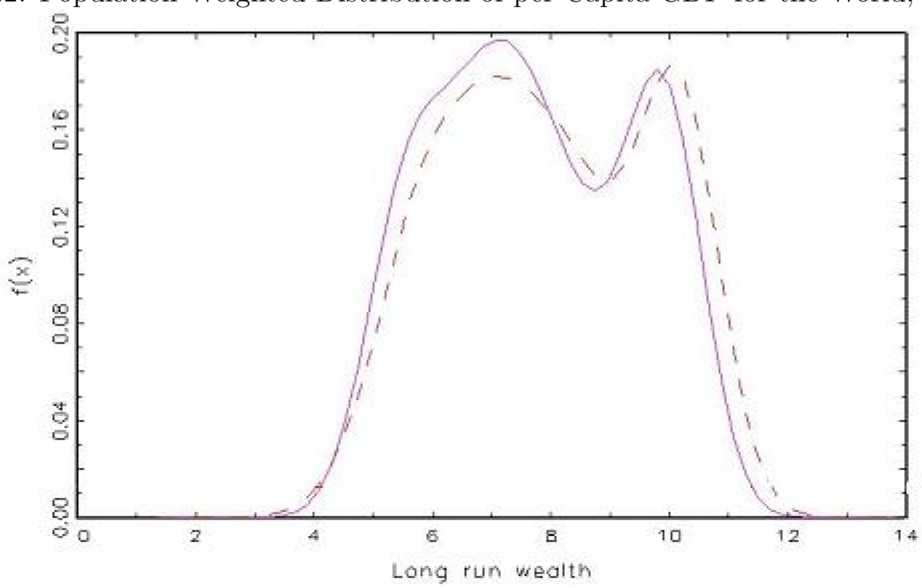


Table A.1: Unadjusted Summary Statistics, 1990-2005

	Unadjusted					Population Weighted		
Year	Means	Medians	Std. Dev.	Maximum	Minimum	Mean	Median	Std. Dev.
	Mixture							
	GDP Per Capita							
1990	7.40	7.17	1.64	10.41	4.80	7.26	10.41	1.80
1995	7.43	7.25	1.71	10.51	4.03	7.30	10.47	1.78
2000	7.52	7.19	1.73	10.74	4.45	7.36	10.52	1.76
2005	7.62	7.35	1.73	10.86	4.51	7.47	10.60	1.70
	Life Expectancy							
1990	63.10	65.61	11.37	78.84	31.17	63.85	78.84	9.47
1995	63.73	67.53	12.06	79.54	31.69	64.90	78.74	9.79
2000	64.19	69.05	12.97	81.08	37.90	65.69	80.88	10.52
2005	65.10	70.39	13.48	82.08	34.97	66.48	81.58	10.75
	Wealth							
1990	245.85	239.89	72.10	375.67	122.44	242.68	207.47	74.61
1995	247.93	245.36	75.16	378.98	107.08	245.72	222.59	74.12
2000	251.32	244.92	77.09	385.90	120.06	248.81	215.21	74.40
2005	255.89	249.47	77.79	392.49	124.19	253.42	215.17	72.83
	Africa							
	GDP per Capita							
1990	5.99	5.79	0.88	8.31	4.80	5.87	8.31	0.85
1995	5.89	5.71	0.93	8.31	4.03	5.77	8.31	0.88
2000	5.95	5.77	0.92	8.26	4.45	5.81	8.26	0.90
2005	6.04	5.86	0.95	8.44	4.51	5.91	8.26	0.91
	Life Expectancy							
1990	50.60	50.75	7.50	64.46	31.17	49.95	64.46	6.88
1995	49.75	50.06	7.41	66.90	31.69	48.72	66.90	6.85
2000	48.45	46.83	7.03	68.81	37.90	47.21	68.81	6.63
2005	48.66	46.93	7.78	70.38	34.97	47.82	63.49	6.85
	Rest							
	GDP Per Capita							
1990	8.11	7.96	1.47	10.41	5.17	7.50	10.41	1.82
1995	8.20	8.10	1.47	10.51	5.30	7.58	10.51	1.76
2000	8.31	8.19	1.48	10.74	5.41	7.67	10.52	1.73
2005	8.42	8.39	1.47	10.86	5.45	7.79	10.86	1.66
	Life Expectancy							
1990	69.34	70.62	6.95	78.84	48.92	66.27	78.84	7.60
1995	70.71	71.46	6.60	79.54	49.02	67.87	79.54	6.93
2000	72.07	72.48	6.39	81.08	50.68	69.25	80.88	6.67
2005	73.31	73.17	6.24	82.08	52.61	70.26	82.08	6.65