Making Inferences About Rich Country-Poor Country Convergence: The Polarization Trapezoid and Overlap Measures

Gordon Anderson* Oliver Linton†
University of Toronto The London School of Economics

Teng Wah Leo‡
St. Francis Xavier University

30 September, 2009

*Department of Economics, University of Toronto. Email Address: anderson@chass.utoronto.ca
†Department of Economics, London School of Economics. Email Address: o.linton@lse.ac.uk
‡Department of Economics, St. Francis Xavier University. Email Address: tleo@stfx.ca
Abstract

Underlying the unresolved debate over whether the gap between rich and poor country GNP per capita has narrowed is a concern for wellbeing. The issue is really about the changing shapes of distributions of wellbeing indicators. As limiting cases convergence between rich and poor country groups can be brought about by countries within groups becoming less alike without any diminution of growth rate differentials between them or it can be brought about by reductions in these differentials without any diminution of within group identity. In essence the debate is about the extent to which rich and poor countries are polarizing, a subject first theoretically explored by Esteban and Ray (1994). The empirical issue is about whether separate groups can be identified in the overall distribution and whether they are tending toward common or distinct equilibria. This paper proposes two simple statistics for the problem, the Overlap measure and the Trapezoidal measure, changes in which reflect a combination of increasing (decreasing) subgroup location differences and decreasing (increasing) subgroup spreads which are the characteristics of polarization (convergence). The former statistic is of use when the sub-distributions are identified, while the latter can be used whether or not the subgroups are identified. These techniques are applied to the examination of convergence in GDP per capita between rich and poor nations when growth is viewed either as a wellbeing index or a technology index (i.e. the data are, or are not, population weighted). It turns out that such a distinction matters, viewed technologically there is divergence, viewed in a wellbeing sense there is convergence. As a collection of countries Africa is diverging from the rest of the world whatever the perspective of growth.
1 Introduction

There has been much debate over whether the gap between rich and poor countries’ GNP 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). Underlying this interest is a concern for wellbeing (and the lack of progress of the poor countries) which judges too much inequality in per capita GNP as a bad thing. The argument has transcended the use of simple per capita GNP measures extending the debate to broader measures that incorporate length and quality of life in the calculus (Decancq, Decoster, and Schokkaert (2009); Becker, Philipson, and Soares (2005)) and to an individualistic sense of income (i Martin 2006). Often this debate has been pursued in terms of the nature of and change in inequality between countries when they are not separately identified as members of rich and poor groups which is a slightly different matter.

Following a growth regression literature which focused on Beta convergence (related to the coefficient on lagged income in a growth regression) and Sigma convergence (related to the conditional variance of incomes) culminating in Barro and Sala-I-Martin (1992) and Galor (1996), there has been extensive interest in examining the relative merits of the Absolute Convergence Hypothesis versus the Club Convergence Hypothesis (see for example Quah (1997), Jones (1997), Paapaa and van Dijk (1998), Bianchi (1997), Durlauf and Quah (1999), Johnson (2000), Islam (2003), Anderson (2004a), Beaudry, Collard, and Green (2005), and Pittau and Zelli (2006)). The latter hypothesis corresponds to a tendency toward multiple modes 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 in that distribution.

The issue is really about the changing nature of the anatomy of distributions of well-being indicators. As limiting cases, convergence between rich and poor groups can be brought about by diminishing within group identity (agents within groups becoming less alike) without any diminution of growth rate differentials between groups or it can be brought about by reductions in these differentials without any diminution of within group identity. This is very much the stuff of a 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 separates itself from pure notions of inequality since it can be readily shown that increased polarization
can either reduce or increase inequality as conventionally measured.

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 per capita GNP 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 per capita GNP nexus is viewed as a technological relationship 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.

Here these issues are addressed by employing new measures of convergence divergence developed for a related literature on polarization. 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 GNP per capita and life expectancy of 123 countries over the period 1990-2005 drawn from the World Bank 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 un-weighted results. For this reason they have been excluded from the analysis.

In the following, section 1 considers the links between the Convergence and Polarization literatures. Section 2 introduces the new measures and outlines their statistical properties. The application is reported in section 3 and conclusions are drawn in section 4.

\footnote{In fact their inclusion does not alter the substantive results at all.}
2 Convergence and Polarization

The issue is very much about whether or not separate groups or clubs can be identified in the overall distribution and whether they are tending toward common (converging) or distinct (diverging) equilibria or, put another way, the question is whether or not the groups are depolarizing or polarizing in a particular fashion. Examining whether or not the poor nation - rich nation divide is diminishing in this context is really about eliciting from an observed mixture of distributions how the sub-distributions (representing the respective clubs) are behaving in terms of their movement or separation. In the convergence literature, polarization (divergence) has been inappropriately associated with non-decreasing variance of the overall mixture distribution and convergence (depolarization) associated with its non-increasing variance but the polarization literature has been at pains to distinguish itself from pure inequality measurement. 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 or members of respective clubs becoming more alike) and increased between group alienation (usually associated with increasing between group differences in location or members of different clubs becoming more un-alike). It has nothing to do with trends in the global variance which is a monotonic increasing function of absolute between group location differences and within group dispersions 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 subgroup or club distributions, and the mixture distribution becomes

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

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 subgroup variances (more homogeneous subgroup behavior) will be accompanied by diminishing variance in the population mixture (contrary to the Club Convergence hypothesis). This is essentially polarization brought about by increasing
within sub-group identity or association rather than increased alienation between groups (see Duclos, Esteban, and Ray (2004)) and is very much in the nature of within club convergence and between club divergence. Of course the reverse process will yield a trend toward uni-modality 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 subgroup or global variance.

Examination of the Absolute versus Club Convergence Hypotheses thus boils down to whether or not the sub-distributions are moving toward each other or whether they are moving apart. Observed changes in the variance of the overall mixture distribution is a misleading statistic for this purpose, 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 is a cumbersome approach and a much more simplistic and easier to understand statistical indicator is required. This paper proposes two simple statistics for the problem, the Overlap measure and the Trapezoidal measure, changes in which reflect a combination of increasing (decreasing) subgroup location differences and decreasing (increasing) subgroup spreads which are the characteristics of polarization (convergence). The former statistic is only of use when the sub-distributions are identified, the latter can be used whether or not the subgroups are identified.

3 The Method

Suppose the rich and poor distributions are separately identified and let $x_{m,p}$ be the value of log GDP per capita ($x$) 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. 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). However when the distributions are not separately identified but are embedded in a mixture, the overlap measure is no longer available, while fortunately the Trapezoid is, provided the mixture is bimodal (See figure [1]). 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 this work.

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

\[
OV = \int_{-\infty}^{\infty} \min\{f_p(x), f_r(x)\} \, dx
\]

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

\[
\begin{align*}
C_{f_p,f_r} &= \{x \in \mathbb{R} : f_p(x) = f_r(x) > 0\}; \quad p_0 = \Pr(X \in C_{f_p,f_r}) \\
C_{f_p} &= \{x : f_p(x) < f_r(x)\}; \quad p_p = \Pr(X \in C_{f_p}) \\
C_{f_r} &= \{x : f_p(x) > f_r(x)\}; \quad p_r = \Pr(X \in C_{f_r})
\end{align*}
\]
The kernel estimator of \( \theta = \int \min \{ f_p(x), f_r(x) \} \, dx \) is shown to be normally distributed of the form:

\[
\sqrt{n}(\hat{\theta} - \theta) - a_n \Rightarrow N(0, v)
\]

where

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

where \( a_n \) and \( \sigma_0^2 \) are bias correction factors (see Anderson, Linton, and Whang (2009) for details).

With respect to polarization, the intensity of within group association 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 p.d.f. be. That the Euclidean distance between the two modal points represents the sense of alienation between the two groups is somewhat more obvious. 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, 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’m 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 non-poor distributions are separately identified in \( J \) dimensions the indicator \( BIPOL \) may be written as:

\[
BIPOL = \frac{1}{2} \left\{ f_p(x_{m,p}) + f_r(x_{m,r}) \right\} \frac{1}{\sqrt{J}} \left\{ \sum_{j=1}^{J} \frac{(x_{m,p,j} - x_{m,r,j})^2}{\mu_j} \right\}^{\frac{1}{2}}
\] (2)

When the groups are not separately identified \( (NI) \) and the index is calculated from the modal points of the mixture distribution, noting that the poor and rich modes may be written in terms of the underlying distributions as:

\[
f(x_{m,p}) = f_r(x_{m,r}) + \omega (f_p(x_{m,p}) - f_r(x_{m,r}))
\] (3)

\[
f(x_{m,r}) = f_p(x_{m,p}) + \omega (f_r(x_{m,r}) - f_p(x_{m,p}))
\] (4)

The index may also be written as:

\[
BIPOL_{NI} = \frac{1}{2} \left\{ f(x_{m,p}) + f(x_{m,r}) \right\} \left\{ \frac{1}{\sqrt{J}} \left\{ \sum_{j=1}^{J} \frac{(x_{m,p,j} - x_{m,r,j})^2}{\mu_j} \right\}^{\frac{1}{2}} \right\}^{\frac{1}{2}}
\] (5)
The estimator of the trapezoid is given by:

\[
\hat{BIPOL} = \frac{1}{2} \left( \hat{f}(\widehat{x}_{m,p}) + \hat{f}(\widehat{x}_{m,r}) \right) (\widehat{x}_{m,p} - \widehat{x}_{m,r})
\] (6)

Here \(x_{m,i}\) is the mode for group \(i\) and \(f(x_{m,i})\) is the value of the p.d.f. 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 \(\hat{BIPOL}\) as:

\[
\frac{(nh^3)^{1/2}(\hat{BIPOL} - BIPOL)}{\text{Bias}, \frac{1}{4} \left\{ f_i(x_{m,i}) + f_j(x_{m,j}) \right\}^2 \left\{ \frac{f_i(x_{m,i})}{|f_i''(x_{m,i})|^2} + \frac{f_j(x_{m,j})}{|f_j''(x_{m,j})|^2} \right\} \|K'\|_2^2} \xrightarrow{D} N(\text{Bias}, \frac{1}{4} \left\{ f_i(x_{m,i}) + f_j(x_{m,j}) \right\}^2 \left\{ \frac{f_i(x_{m,i})}{|f_i''(x_{m,i})|^2} + \frac{f_j(x_{m,j})}{|f_j''(x_{m,j})|^2} \right\} \|K'\|_2^2)
\] (7)

Tests are based on the trapezoid measure being asymptotically normally distributed with a variance approximately equal to

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

where \(x_{m,j}, j = \{r, p\}\) are the modes of the respective distributions, \(K\) 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 \(f(\cdot)\) can be estimated (again based upon a Gaussian kernel) as:

\[
f^s(x) = \frac{(-1)^2}{nh^{s+1}} \sum_{i=1}^n K^2 \left( \frac{x_i - x}{h} \right)
\] (8)

where for \(s = 2\),

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

When the poor and rich distributions are not identified life gets a little more complicated but the principles are the same. The mixture distribution is always observed, the question is whether it is possible to identify the sub-distributions in the mixture (Or at least can the locations and the heights of the sub-distribution peaks be identified)? In the application to be reported in the following section, this has not presented a problem, however it is not always so simple. Some discussion of modality detection is contained in a “Bump Hunting” literature reported in Silverman (1986), 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.
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 and the Rest of the World as separate entities.

In terms of representing wellbeing, the use of log GDP per capita involves two major issues. Firstly, growth regressions have very much a flavour of representative agent models with country i’s log(GDP per capita) being the log consumption (or income) of the representative agent of the i-th country. When used in un-weighted growth regressions the agent from Ireland (3.5 million population in 1990) has exactly the same weight as the people of China (1135 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 \( \left( \int_0^T U(C(t))e^{-rt} dt \right) \) subject to the present value of lifetime wealth \( \left( \int_0^T Y(t)e^{-rt} dt \right) \) where \( U(\cdot) \) is an instantaneous felicity function, \( Y \) is income, \( r^* \) is the individuals rate of time preference and \( r \) is the market lending rate. Browning and Lusardi (1996) showed that this taken together with the assumption of a constant relative risk aversion and no bequest motive preference structure leads to a consumption smoothing model of the form:

\[
C(t) = e^{\left( \frac{(r-r^*)t}{\zeta} \right)} C(0) \tag{10}
\]

where \( \zeta \) is the risk aversion coefficient and by implication \( g = (r-r^*)/\zeta \) is the consumption growth rate\(^2\). The point is that the wellbeing of the representative agent very clearly depends upon her life expectancy and since life expectancy varies considerably across

\(^2\)The empirical counterpart of this equation is the familiar random walk model:

\[ \ln C(t) = \ln C(t-1) + g + \epsilon(t) \]

which is the basis of the cross country convergence regressions familiar in the growth literature.
countries (i.e. agents) it needs to be accommodated in the calculus. Here wellbeing will be considered to be represented by:

\[
\int_0^T \ln(\text{GDP p.c.}) e^{-gt} dt = \frac{\ln(\text{GDP p.c.})}{-g} (e^{-gt} - 1)
\]  

(11)

which is essentially an aggregation (discounted present value) of instantaneous utilities represented by \(\ln(\text{GDP per capita})\) where the discount rate was set equal to the growth rate. Growth rates used were the average growth rate over the sample unless it was non-positive in which case 0.0067% was used. 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 convergence made using multivariate versions of the overlap and trapezoidal measures.

Table A.1 in the appendix reports the summary statistics of the data and standard normal tests of the changes are presented in table 1 over the sample period. From these tables it may be seen that there has been significant growth in per capita GNP and increases in life expectancy over the observation period in the full sample whether population weighted or not. The variances have generally increased (but not to a substantive extent) lending some support to the divergence hypothesis. The results for the African nations are not so clear cut with all changes 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 un-weighted sample but rose in the weighted sample, while the results for the non-African countries reflect those of the full sample.

Kernel estimates of the univariate mixture distributions (unweighted and population weighted respectively) of lifetime wellbeing are reported in table 2 and depicted in figures 2 and 3. It is immediately apparent that population weighting makes a considerable difference to the distribution’s shape, emphasizing somewhat 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.

---

3It is of interest to see how different are the pure GDP per capita distributions which are depicted in appendix.
Figure 2: Unweighted Mixture Distribution, 1990-2005

Figure 3: Population Weighted Mixture Distributions, 1990-2005
Table 1: Difference in Means Tests, GNP Life Expectancy, 1990-2005

<table>
<thead>
<tr>
<th>Population</th>
<th>GNP per capita</th>
<th>Life Expectancy</th>
<th>N</th>
<th>t</th>
<th>Pr(T &gt; t)</th>
<th>t</th>
<th>Pr(T &gt; t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>9.2122</td>
<td>-3.8621</td>
<td>123</td>
<td>0.0000</td>
<td>0.0000</td>
<td>1.486</td>
<td>0.9314</td>
</tr>
<tr>
<td>Africa</td>
<td>-1.0587</td>
<td>1.486</td>
<td>41</td>
<td>0.1449</td>
<td>0.9314</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The Rest</td>
<td>-13.0865</td>
<td>-18.1695</td>
<td>82</td>
<td>0.0000</td>
<td>0.0000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Population</th>
<th>Weighted</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>-7.5185</td>
<td>-7.5185</td>
<td>123</td>
<td>0.0000</td>
<td>0.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Africa</td>
<td>-1.6787</td>
<td>-0.9692</td>
<td>41</td>
<td>0.0466</td>
<td>0.1662</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The Rest</td>
<td>-12.7969</td>
<td>-12.7969</td>
<td>82</td>
<td>0.0000</td>
<td>0.0000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The trapezoidal tests reported in table 2 highlight the impact of population weighting, with the un-weighted sample strongly rejecting the null of convergence and the weighted sample strongly supporting the hypothesis of convergence. This no doubt reflects the Martin (2006) finding of convergence when global inequality is treated in an individualistic sense as opposed to when it is addressed in a between country (i.e. population un-weighted) 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).

Table 2: Mixture Trapezoids

<table>
<thead>
<tr>
<th></th>
<th>Unweighted Sample</th>
<th>Weighted Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1990</td>
<td>2005</td>
</tr>
<tr>
<td>Trapezoid Value</td>
<td>0.6668</td>
<td>0.7041</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.00044</td>
<td>0.00037</td>
</tr>
<tr>
<td>t Statistic for Difference</td>
<td>64.9287</td>
<td>-52.1771</td>
</tr>
</tbody>
</table>

Turning to a comparison of Africa and the Rest of the World, figures 4 and 5 depicts the kernel estimates of the unweighted and population weighted distributions respectively, where again population weighting is seen to have a considerable impact on the shapes of the distributions though this time it does not appear to affect the convergence results.
reported in table 3. Note here with respect to the weighted distributions, Africa has significantly polarized with respect to both the poor and rich groups in the rest of the world, whereas the poor in the Rest of the World have closed the gap with the rich (remember China and India have been excluded from this analysis). The latter observation is evident from figure 5 with the increased density at the mode and reduced distance between modes between 1990 and 2005. The trapezoid and overlap measures of table 3 both strongly reject the hypothesis of convergence of Africa and the Rest of the World reflecting increases in the trapezoidal area and reductions in the overlap for both population weighted and un-weighted calculations.

The robustness of the univariate result for convergence in income and life expectancy can be addressed in a multivariate framework, treating GDP per capita and life expectancy as separate variables as opposed to the lifetime wellbeing measure of equation 11, using both trapezoidal and overlap measures. The results are reported in table 4 and the significant increase in the area of the trapezoid and fall in the overlap measure confirm the results in the single variable framework.

Table 3: Africa and the Rest Comparisons

<table>
<thead>
<tr>
<th></th>
<th>Unweighted Distributions</th>
<th>Weighted Distributions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1990</td>
<td>2005</td>
</tr>
<tr>
<td>Trapezoid Value</td>
<td>0.8626</td>
<td>0.9043</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.0032</td>
<td>0.0024</td>
</tr>
<tr>
<td>t Statistic for Difference</td>
<td>-8.8740</td>
<td>-122.1292</td>
</tr>
<tr>
<td>Trapezoid Value</td>
<td>1.6994</td>
<td>1.8401</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.00257</td>
<td>0.00195</td>
</tr>
<tr>
<td>t Statistic for Difference</td>
<td>-43.6418</td>
<td>-72.6587</td>
</tr>
<tr>
<td>Overlap Measure</td>
<td>0.2632</td>
<td>0.1226</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.05110</td>
<td>0.03737</td>
</tr>
<tr>
<td>t Statistic for Difference</td>
<td>2.2205</td>
<td>3.3144</td>
</tr>
</tbody>
</table>
Figure 4: Distribution for Africa and the Rest, 1990-2005

Figure 5: Population Weighted Distribution for Africa and the Rest, 1990-2005
<table>
<thead>
<tr>
<th></th>
<th>Unweighted Distributions</th>
<th>Weighted Distributions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1990</td>
<td>2005</td>
</tr>
<tr>
<td>Trapezoid Value</td>
<td>1.5284</td>
<td>1.7921</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.1225</td>
<td>0.0992</td>
</tr>
<tr>
<td>t Statistic for Difference</td>
<td>-1.6731</td>
<td></td>
</tr>
<tr>
<td>Overlap Measure</td>
<td>0.2516</td>
<td>0.1298</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.0488</td>
<td>0.0241</td>
</tr>
<tr>
<td>t Statistic for Difference</td>
<td></td>
<td>2.2365</td>
</tr>
</tbody>
</table>

### 5 Conclusion

Convergence is about the changing nature of the anatomy of distributions of wellbeing indicators. As limiting cases, separately identified rich and poor club convergence can be brought about by diminishing within club identity (agents within clubs becoming less alike) without any diminution of club growth rate differentials (club locations converging) or it can be brought about by club locations converging (diminishing between club alienation) without any diminution of within club identity. As limiting cases, diminishing within club identity increases global variance whereas diminishing between club alienation reduces global variance rendering trends in global inequality invalid as instruments for identifying trends in global convergence. On the other hand, measures of the extent to which distributions of wellbeing indicators overlap or measures which are monotonically increasing functions of the extent of a distribution’s modality and the extent to which their modal coordinates differ provide very reliable instruments for identifying trends in global convergence. A second issue in the convergence calculus is to decide whether the concern is convergence in individualistic wellbeing or convergence in wellbeing producing technologies. If the former, an inter-country analysis requires consideration of population weighting issues, while the latter does not. A third issue is to consider convergence in terms of lifetime wellbeing based upon some combination of measures of annual expected income and life expectancy.

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 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 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 club-rich club convergence in the world wellbeing distribution when considered in an individualistic basis, there is 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.
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 $K$ independent $N(0,1)$ is $1/\sqrt{K}$ times the height of one $N(0,1)$.

The estimator of the trapezoid:

$$BIPOL = \frac{1}{2} \left( \hat{f}_p(\hat{x}_{m,p}) + \hat{f}_r(\hat{x}_{m,r}) \right) \frac{1}{\sqrt{J}} \sqrt{\sum_{j=1}^{J} (\hat{x}_{m,p,j} - \hat{x}_{m,r,j})^2 \mu_j}$$

(A-1)

where $\mu_j$ is the average of the modes in the $j$’th dimension and where $x_{m,i}$ 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 $\hat{f}_i, i \in \{p, r\}$ refer to the kernel estimates of the population density function.

Considering first the one-dimensional version,

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

(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),

$$(nh^3)^{1/2}(\hat{x}_{m,i} - x_{m,i}) \overset{D}{\longrightarrow} N \left( \text{Bias}, \frac{f_i(x_{m,i})}{[\hat{f}_i(x_{m,i})]^2} ||K||^2_2 \right)$$

(A-3)

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

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

(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,

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

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

\[
Z_n(t) \Rightarrow Z(t) = \frac{f''(x_{m,i}) t^2}{2} + (-1)^{q+1} \frac{f^{(q+1)}(x_{m,i})}{q!} dB_q t + Y t
\]

where \( Y \) is a normally distributed random variable, \( N(0, f(x_{m,i}) ||K'||_2^2) \), \( q \geq 2 \) is an integer, \( \lim_{n \to \infty} (nh^3 + 2q)\frac{1}{2} = d \), \( f^{(q+1)} \) 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,

\[
b^{-2} \left[ \hat{f}_i(x_{m,i}) - \hat{f}_i(x_{m,i}) \right] = Z_n(t) \Rightarrow Z(\hat{t})
\]  
(A-6)

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

\[
(-1)^q \frac{d}{q!} \frac{f^{(q+1)}(x_{m,i})}{f''(x_{m,i})} B_q
\]

Therefore,

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

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

\[
\hat{f}_i(x_{m,i}) - f_i(x_{m,i}) = o_p(1)
\]  
(A-8)

So that equation (A-4) is,

\[
\hat{f}_i(x_{m,i}) - f_i(x_{m,i}) = O_p(n^{-1}h^{-3})
\]  
(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,

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

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

\[
\sum_{i \in \{r,p\}} (-1)^q \frac{d}{q!} \frac{f^{(q+1)}(x_{m,i})}{f''(x_{m,i})} B_q
\]
It follows that in this regular case,

\[
(nh^3)^{1/2}(\overline{\text{BIPOL}} - \text{BIPOL}) \xrightarrow{D} N \left( \text{Bias}, \frac{1}{4} \left\{ f_i(x_{m,i}) + f_j(x_{m,j}) \right\}^{2} \left\{ \frac{f_i(x_{m,i})}{|f_i'(x_{m,i})|^2} + \frac{f_j(x_{m,j})}{|f_j'(x_{m,j})|^2} \right\} \|K'\|^2 \right)
\]  

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

<table>
<thead>
<tr>
<th>Year</th>
<th>Means</th>
<th>Medians</th>
<th>Std. Dev.</th>
<th>Maximum</th>
<th>Minimum</th>
<th>Mean</th>
<th>Median</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unadjusted</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>GDP Per Capita</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>7.4037</td>
<td>7.1684</td>
<td>1.6427</td>
<td>10.4127</td>
<td>4.8035</td>
<td>7.2567</td>
<td>10.4127</td>
<td>1.8010</td>
</tr>
<tr>
<td></td>
<td>Life Expectancy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>63.0950</td>
<td>65.6075</td>
<td>11.3704</td>
<td>78.8368</td>
<td>31.1730</td>
<td>63.8509</td>
<td>78.8368</td>
<td>9.4735</td>
</tr>
<tr>
<td>1995</td>
<td>63.7255</td>
<td>67.5307</td>
<td>12.0584</td>
<td>79.5363</td>
<td>31.6940</td>
<td>64.9047</td>
<td>78.7405</td>
<td>9.7903</td>
</tr>
<tr>
<td>2000</td>
<td>64.1923</td>
<td>69.0512</td>
<td>12.9745</td>
<td>81.0761</td>
<td>37.9048</td>
<td>65.6878</td>
<td>80.8780</td>
<td>10.5179</td>
</tr>
<tr>
<td>2005</td>
<td>65.0975</td>
<td>70.3944</td>
<td>13.4848</td>
<td>82.0754</td>
<td>34.9659</td>
<td>66.4818</td>
<td>81.5805</td>
<td>10.7484</td>
</tr>
<tr>
<td></td>
<td>Africa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>GDP per Capita</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>5.9901</td>
<td>5.7853</td>
<td>0.8762</td>
<td>8.3134</td>
<td>4.8035</td>
<td>5.8664</td>
<td>8.3134</td>
<td>0.8541</td>
</tr>
<tr>
<td>1995</td>
<td>5.8905</td>
<td>5.7135</td>
<td>0.9319</td>
<td>8.3086</td>
<td>4.0346</td>
<td>5.7742</td>
<td>8.3086</td>
<td>0.8785</td>
</tr>
<tr>
<td>2000</td>
<td>5.9470</td>
<td>5.7679</td>
<td>0.9222</td>
<td>8.2628</td>
<td>4.4547</td>
<td>5.8070</td>
<td>8.2628</td>
<td>0.9041</td>
</tr>
<tr>
<td>2005</td>
<td>6.0355</td>
<td>5.8593</td>
<td>0.9465</td>
<td>8.4443</td>
<td>4.5106</td>
<td>5.9140</td>
<td>8.2646</td>
<td>0.9085</td>
</tr>
<tr>
<td></td>
<td>Life Expectancy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>50.5953</td>
<td>50.7493</td>
<td>7.4988</td>
<td>64.4622</td>
<td>31.1730</td>
<td>49.9515</td>
<td>64.4622</td>
<td>6.8791</td>
</tr>
<tr>
<td>2000</td>
<td>48.4466</td>
<td>46.8340</td>
<td>7.0348</td>
<td>68.8051</td>
<td>37.9048</td>
<td>47.2108</td>
<td>68.8051</td>
<td>6.6346</td>
</tr>
<tr>
<td>2005</td>
<td>48.6648</td>
<td>46.9278</td>
<td>7.7794</td>
<td>70.3757</td>
<td>34.9659</td>
<td>47.8232</td>
<td>63.4868</td>
<td>6.8545</td>
</tr>
<tr>
<td></td>
<td>Rest</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>GDP Per Capita</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Life Expectancy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>69.3448</td>
<td>70.6177</td>
<td>6.9521</td>
<td>78.8368</td>
<td>48.9161</td>
<td>66.2702</td>
<td>78.8368</td>
<td>7.5960</td>
</tr>
<tr>
<td>1995</td>
<td>70.7136</td>
<td>71.4604</td>
<td>6.5990</td>
<td>79.5363</td>
<td>49.0177</td>
<td>67.8658</td>
<td>79.5363</td>
<td>6.9268</td>
</tr>
<tr>
<td>2000</td>
<td>72.0651</td>
<td>72.4775</td>
<td>6.3929</td>
<td>81.0761</td>
<td>50.6708</td>
<td>69.2538</td>
<td>80.8780</td>
<td>6.6749</td>
</tr>
<tr>
<td>2005</td>
<td>73.3138</td>
<td>73.1717</td>
<td>6.2426</td>
<td>82.0754</td>
<td>52.6143</td>
<td>70.2607</td>
<td>82.0754</td>
<td>6.6531</td>
</tr>
</tbody>
</table>
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
References


