Boats and Tides and "Trickle Down" Theories:
What Stochastic Process Theory has to say about Modeling Poverty, Inequality and Polarization.

Gordon Anderson

University of Toronto
Economics Department

This paper looks at the implications for empirical wellbeing analysis of conventional assumptions usually made by macroeconomists concerning the processes that underlay income and consumption outcomes. Various forms of poverty, inequality and income mobility structures are considered and much of the conventional wisdom concerning the implications for wellbeing of economic growth is questioned. The results are applied to the distribution of GDP in the continent of Africa.

Introduction.

The aphorism that "a rising tide raises all boats" and the theory that advances in economic well-being of the rich ultimately trickle down to the poor have frequently been cited as reasons for believing that growth will elevate the poor from poverty. These are essentially notions about the nature of income or consumption processes as stochastic processes. Economists interested in growth, consumption and convergence issues of various forms have a long tradition of modeling income or consumption as a stochastic process (usually of the random walk variety). This is presumably because such processes provide a good description of the progress of income and consumption paths for modeling purposes but also because such formulations, in the form of growth regressions, provided a useful way of relating growth rates to initial conditions. Surprisingly, for the link does not appear to have been made very often in the income size distribution and economic well being literatures [1], such models have implications for, and provide predictions as to, the progress of inequality or poverty that would be of interest to those interested in various aspects of empirical well-being. Also somewhat surprisingly these predictions do not always accord with conventional wisdom underlying "trickle down" theories and "tides and boats" aphorisms.

Stochastic process theory also provides a motivation for fitting particular size distributions of income or consumption (since the nature of the stochastic process has implications for the nature of the size distribution of income). There are many advantages

---

[1] (Blundell and Lewbel (2008), Deaton and Paxson (1994), Meghir and Pistaferri (2004), Meyer and Sullivan (2003) and O'Neill (2005), are exceptions)

associated with fitting size distributions. Poverty calculations of the non-parametric variety can be difficult, especially when the poverty group is small in number relative to the size of the population, since information on the relevant tail of the distribution is sparse and changes in the nature of that tail can be very difficult to get a handle on (see Davidson and Duclos (2008) for a discussion that highlights this problem). Furthermore if incomes are truly governed by such processes anti-poverty (inequality) policies need to focus on changing the nature of the process or at least mitigating its effects and fully understanding the nature and implications of the processes will help in this regard. Indeed defining a poverty frontier or lower bound below which incomes are not permitted to fall, such as a social security net, as will be seen, can become part of the process, changing its nature and the nature of the resultant size distribution of income. This in turn provides a test of the effectiveness of the policy in terms of the extent to which the nature of the distribution has changed.

Alternatively one may wish to define the poor as a sub group, an entity in itself, with a unique stochastic process defining its path as opposed to the paths of other groups in society. The societal income distribution then in effect becomes a mixture distribution governed by the variety of processes defining the separate classes and the mixture coefficients which define the memberships of those classes. Anti poverty policies can then focus on the changing the nature of these processes relative to those of other groups in society and ideas from the convergence literature become relevant. In these circumstances the way poverty is measured also needs to be reviewed since poverty is

now about membership of a class and the way the stochastic process describing that class proceeds.

Two early front runner's for describing the size distribution of income or consumption were the Pareto distribution and the Lognormal distribution[2], subsequently it has been learned that they are linked via stochastic process theory. Pareto (1897) felt that his distribution was a law which governed the size distribution of incomes, Gibrat (1931), working with firm sizes, used central limit theorem type arguments to demonstrate that a sequence of successive independent proportionate "close to one" shocks to an initial level of a variable would yield an income the log of which was normally distributed regardless of the distribution governing the shocks. Kalecki (1945) showed that Gibrat's result could be obtained from a stationary process as well. Gabaix (1997), working with city size distributions, highlighted the link in showing that if a process such as that proposed by Gibrat (or Kalecki) was subjected to a reflective lower boundary, bouncing back the variable should it hit the boundary from above, the resulting distribution would be Pareto[3]. Obviously a social security net of some kind, such as a legislated low income cut-off below which no one was permitted to fall, would constitute such a boundary[4] for an income process. Both of these notions regarding the shape of income size distributions draw on theories of stochastic processes which, if empirically verified, will also tell us much about the progress of poverty and or inequality however defined.

---

[2] Conventional wisdom was that Pareto fit well in the tails whereas the lognormal fit well in the middle, (Harrison (1984), Johnson, Kotz and Balakrishnan (1994)).

[3] This has been established before, Harrison (1987) and Champernowne (1953) demonstrate a somewhat similar result. Reed (2006) provides an alternative link between the Lognormal and Pareto and provides rationales from stochastic process theory for more complex size distributions.

[4] Both Canadian and British governments have vowed to eliminate child poverty. The Millenium goals and $1 and $2 poverty frontiers may also be construed as such potential frontiers.

Before examining what such models imply for the progress of poverty or inequality one needs to be clear as to what sort of poverty or inequality it is that is in question. There has been considerable debate about the nature of poverty measurement as to whether it should be an absolute or a relative measure. The issue entertained the minds of the founders of the discipline. Adam Smith (1776) can be interpreted to have had a relative view of poverty viz: "..By necessaries I understand, not only the commodities which are indispensably necessary for the support of life, but whatever the custom of the country renders it indecent for creditable people, even the lowest order, to be without." Similarly Ferguson (1767) states "The necessary of life is a vague and relative term: it is one thing in the opinion of the savage; another in that of the polished citizen: it has a reference to the fancy and to the habits of living".Townsend (1985) (the major advocate of the relative measure in recent times) and Sen (1983) (who favours a basic needs formulation) have lead the debate regarding the two approaches, both claim consistency with the intent of Smith's thoughts largely via different interpretations of the words decent, creditable etc. Interestingly enough no such debate seems to have taken place regarding inequality, though invariably relative inequality measures (Gini and Shutz coefficients for example) seem to have been favored. Some absolute measures (variance levels and quantile differences for example) have currency and recently notions of polarization which is related to, but not the same as, inequality have gained favour (see Duclos, Esteban and Ray (2004) for details).

Here the theoretical implications of stochastic processes for absolute and relative poverty and inequality measures as well as polarization indices will be outlined and the results employed in looking at the stochastic processes underlying the per capita GDP of African nations and considering what they imply for the progress of poverty and inequality on that continent. After a consideration of the implications of some aspects of relatively simple stochastic processes for the progress of poverty and inequality in section 2 more complex structures are considered in section 3. These ideas are considered in the light of data on per capita GDP for African nations over the period 1985 to 2005 in section 4 and conclusions are drawn in section 5.

2. Gibrat's Law, Kalecki's Law, The Pareto Distribution and notions of absolute and relative poverty and inequality.

Starting off with Gibrat's law of proportionate effects in a discrete time paradigm suppose that x_t, the income of the representative agent at period t, follows the law of proportionate effects with δ_t its income growth rate in period t, T the elapsed time period of earnings with x_0 the initial income. Thus:

$$x_t = (1+\delta_{t-1})x_{t-1}; \quad and \quad x_T = x_0 \prod_{i=1}^{T-1}(1+\delta_i) \quad [1]$$

Assuming the δ's to be independent identically distributed random variables with a small (relative to one) mean μ and finite variance σ^2 it may be shown that for an agents life of T

years with starting income x_0 the income size distribution of such agents would be of the form[5]:

$$\ln(x_T) \sim N((\ln(x_0) + T(\mu + 0.5\sigma^2)), T\sigma^2) \quad [2]$$

These types of models are very close to the cross – sectional growth (or Barro) regressions familiar in the growth and convergence literature (see Durlauf et.al.(2005) for details) except that the properties of the error processes they engender are usually ignored in cross-sectional comparisons, in particular the variance of the process is heteroskedastic increasing in a cumulative fashion through time implying increasing absolute inequality. Note that [2] would also be the consequence of a process of the form:

$$\ln(x_t) = \ln(x_{t-1}) + \psi + e_t$$

which had started at t=0 and had run for T periods where e_t was an i.i.d. $N(0,\sigma^2)$ and where $\psi = \mu + 0.5\sigma^2$. Indeed the i.i.d. assumption regarding the δ's is much stronger than needed, under conditions of 3rd moment boundedness, log normality can be established for sequences of non-independent, heteroskedastic and heterogeneous δ (see Gnedenko (1962)). The power of the law, like all central limit theorems, is that a log normal distribution prevails in the limit almost regardless of the underlying distribution of the δ's.

---

[5] The same result can be achieved in the continuous time paradigm by assuming a Geometric Brownian Motion for the x process of the form:
$$dx = \mu x dt + \sigma x dw$$
Where μ is the mean drift σ is a variance factor and dw is the white noise increment of a Weiner process.

Clearly for a needs based (absolute) poverty line (say x*) and growth exceeding $-0.5\sigma^2$ the poverty rate would be 0 in the limit (i.e. $\lim_{T\to\infty} \Phi([(\ln(x^*/x_0)-T(\mu+0.5\sigma^2))/(\sigma\sqrt{T})])$ and for growth less than $-0.5\sigma^2$ the poverty rate would be 1. For a relative poverty line, for example 0.6 of median income (note median income will be $\exp(\ln(x_0)+T(\mu+0.5\sigma^2))$ and the poverty cut-off will be .6 of that value), the poverty rate would be $\Phi([\ln(0.6)/(\sigma\sqrt{T})])$ which obviously increases with time reaching .5 at infinity. The income quantiles in such an income process will not have common trends and, provided growth is sufficiently small[6] , such a society exhibits increasing inequality by most measures that are not location normalized (hereafter referred to as absolute inequality). For aficionados of the Gini what really matters is the growth rate, Lambert (1993) shows that for the Log Normal Distribution with mean and variance θ, γ respectively and with a distribution Function $F(z \mid \theta, \gamma)$ the Gini coefficient may be written in the present context as:

$$2F(\exp(\ln(x_0)+T(\mu+0.5\sigma^2)) \mid \exp(\ln(x_0)+T(\mu+0.5\sigma^2)), T\sigma^2) - 1$$

This will tend to zero as $T \Rightarrow \infty$ when $\mu < -0.5\sigma^2$ and will tend to 1 otherwise, note particularly for zero growth Gini will tend to 1. With respect to relative poverty measures should a "civil society" protect its poor in maintaining its "relative status", for example by defining a poverty cut off such that the poorest 20% of society were considered the poor, then the cut off would exhibit a lower growth rate than mean income. One may thus

---

[6] Many inequality measures are location normalized measures of dispersion, (for example like the Coefficient of Variation and Gini) if the location is increasing slow enough and the dispersion is increasing fast enough inequality by any measure will be increasing

engage propositions such as those mooted in Freidman (2005) by considering the dynamics of the poverty cut – off relative to the mean income.

The Polarization index proposed by Esteban and Ray (1994) may be seen as closely related to the discrete version of the Gini Index since it is of the form:

$$P_\alpha = K \sum_{i=1}^{n} \sum_{j=1}^{n} | x_i - x_j | \pi_i^{1+\alpha} \pi_j$$

where π_i is the probability of being in the i'th cell and $\alpha > 0$ is the index of polarization aversion (when $\alpha = 0$ we have the Gini index) and K is a scale factor (in the Gini it is mean income).

To somewhat muddy the waters Kalecki (1945) generated a lognormal size distribution from a stationary process of the form:

$$\ln x_t - \ln x_{t-1} = \lambda (f(w_t) - \ln x_{t-1}) + e_t \quad [3]$$

With $0 < \lambda < 1$ this corresponds to a partial adjustment model to some equilibrium $f(w_t)$, (which in the context of incomes would be a "fundamentals" notion of long run log incomes). This is essentially a reversion to mean type of process where the mean itself could be a description of the average income level at time t (which incidentally may well be trending through time) but here the variance of the process (and concomitantly absolute inequality) stays constant over time. For $e_t \sim N(0,\sigma^2)$ in the long run $\ln(x_t) \sim N(f(w_t), \sigma^2/\lambda^2)$. There are several observations to be made.

Firstly the pure integrated process story associated with Gibrat's law is not even a necessary condition for lognormality of the income size distribution, such distributions can be obtained from quite different, more generally integrated or non-integrated processes. Secondly stationary processes are in some sense memory-less in that the impacts of the initial value of incomes $f(w_0)$ and the associated shock e_0 disappear after a sufficient lapse of time. On the other hand integrated processes never forget, the marginal impact of the initial size and subsequent shocks remain the same throughout time. Thirdly if $f(w_t)$ were itself an integrated process (if the w's were integrated of order one and f(w) was homogenous of degree one for example) [3] would correspond to an error correction model and incomes would still present as an integrated process in its own right with x and the function of the w's being co-integrated with a co-integration factor of 1. This is the key to distinguishing between "Kalecki's law" and Gibrat's law, the cross-sectional distribution of the former only evolves over time in terms of its mean $f(w_t)$, its variance (written as σ^2 / λ^2) is time independent, whereas the cross distribution of the latter evolves in terms of both its mean and its variance overtime. The distinction has major implications for the progress of poverty and inequality.

Clearly for a needs based (absolute) poverty line (say x*) the poverty rate will depend upon the time profile of $f(w_T)$ in the limit (i.e. $\lim_{T\to\infty} \Phi([(x^*-f(w_T)))/ (\sigma/\lambda)])$ for positive growth it will be 0 and for negative growth it will be 1. For a relative poverty line, 0.6 of median income for example (note median income will be $\exp(f(w_T))$ and the poverty cut-off will be .6 of that value), the poverty rate would be $\Phi([\ln(0.6)/(\sigma/\lambda)])$ which obviously remains constant over time. Inequality measures that are not mean income normalized

will remain constant over time location normalized inequality measures will diminish with positive growth and diminish with negative growth since the Gini coefficient may be written as:

$$2F(\exp(\ln(x_0)+T\mu),\sigma^2 /\lambda^2) - 1$$

which will be 0 for negative growth, 1 for positive growth and constant for zero growth.

Where does Pareto's Law fit in?

Suppose the income process is governed by [1] but now, should x_t fall below x* which is a lower reflective boundary (such as an enforced poverty frontier for example a mandated social security benefit payment), then the process is modified to [1] plus:

$$x_t = x^* \text{ if } (1+\delta_{t-1})x_{t-1} < x^* \quad [1a]$$

Gibrat's Law will no longer hold, in fact after a sufficient period of time the size distribution of x would be Pareto $(F(x) = 1-(x^*/x)^\theta)$ with a shape coefficient $\theta = 1$. In the literature on city size distributions this distribution is known as Zipf's Law and in that literature Gabaix (1999) showed that Zipf's law follows from a Gibrat consistent stochastic process (essentially a random walk) that is subject to a lower reflective boundary. In fact this phenomena, that a random walk with drift that is subject to a lower reflective boundary generates a Pareto distributed variate, has been known in the statistical process literature for some time (see for example Harrison (1987))[7]. In the

---

[7] Champernowne (1953) discovered as much in the context of income size distributions.

present context this has many implications, the Pareto distribution has a very different

shape from the log normal and it would be constant through time, all relative poverty

measures, absolute poverty measures and inequality measures[8] would be constants over

time so that Pareto based predictions provide very powerful tests of the effectiveness of a

mandated social security safety net.

These stochastic theories also have something to say about societal mobility. From a

somewhat different perspective than is usual mobility in a society may be construed as its

agents opportunity for changing rank. Suppose that opportunity is reflected in the chance

that two agents change places and consider two independently sampled agents $x_{it} = x_{it-1} +$

e_{it} and $x_{jt} = x_{jt-1} + e_{jt}$, so that $E(e_{it} - e_{jt}) = 0$ and $V(e_{it} - e_{jt}) = 2\sigma^2$. For the Gibrat model the

probability that agents switch their relative ranks in period t is given by:

$$P(x_{it} > x_{jt} \mid x_{it-1} < x_{jt-1}) \; = \; P(e_{it} - e_{jt} > x_{jt-1} - x_{it-1})$$

By noting that the Gini coefficient is one half the relative mean difference between agents

and that, for the log normal distribution, this may be written as twice the integral of a

standard normal curve over the interval $[0, (\sqrt{(V(x)/2)})]$, the average distance between

two agents $= 4E(x)(\Phi(\sqrt{(V(x)/2)})-\Phi(0))$ and this probability may be written as:

$= \; P((e_{it} - e_{jt})/(\sigma/\sqrt{2}) > 4\exp(\ln(x_0)+T(\mu+0.5\sigma^2))(\Phi(T^{0.5}\sigma/\sqrt{2})-\Phi(0))/(\sigma/\sqrt{2}))$

$= \; P(Z > 4\exp(\ln(x_0)+T(\mu+0.5\sigma^2))(\Phi(T^{0.5}\sigma/\sqrt{2})-\Phi(0))/(\sigma/\sqrt{2}))$

---

[8] The Gini for a Pareto distribution is $1/(2\theta-1)$ which is 1 when the shape coefficient is one because in this case the Pareto distribution has no moments or an infinite mean.

The point is this probability diminishes over time the intuition being that under constant population size the agents are growing further and further apart.

For Kalecki's law note that the independently sampled agents processes may be written as $x_{it} = f(w_t) + (1-\lambda)x_{it-1} + e_{it}$ and $x_{jt} = f(w_t) + (1-\lambda)x_{jt-1} + e_{jt}$, so that $E(e_{it} - e_{jt}) = 0$ and $V(e_{it} - e_{jt}) = 2\sigma^2$, with the inequality being written as:

$$P(x_{it} > x_{jt} \mid x_{it-1} < x_{jt-1}) = P((e_{it} - e_{jt})/(1-\lambda) > x_{jt-1} - x_{it-1})$$

In this case using the Gini relationship to the population mean and the mean difference this probability may be written as:

$$P((1-\lambda)(e_{it} - e_{jt})/(\sigma/\sqrt{2}) > (1-\lambda)4\exp(f(w_t))(\Phi(\sigma/(\lambda\sqrt{2}))-\Phi(0))/(\sigma/\sqrt{2}))$$

$$= P(Z > (1-\lambda)4\exp(f(w_t)) (\Phi(\sigma/((\lambda\sqrt{2}))-\Phi(0))/(\sigma/\sqrt{2}))$$

so that as long as the fundamentals process $f(w_t)$ is constant then so will the mobility in that society. A similar result may be established for Pareto's law. Table 1 summarizes these results:

Table 1.

Wellbeing Type	Type of Stochastic Process		
	Gibrat's Law	Kalecki's Law	Pareto's Law
Absolute Poverty	Increasing or decreasing dependent upon growth rate	Increasing or decreasing dependent upon growth rate	Constant if the reflective boundary is at the poverty cut-off.
Relative Poverty	Increasing with time	Constant	Constant if the reflective boundary is at the poverty cut-off.
Non-Normalized Inequality	Increasing	Constant	Constant
Location	Decreasing with +ve	Decreasing	Constant

Normalized Inequality	growth rate		
Mobility	Diminishing	Constant (if the fundamentals are constant)	Constant (if the fundamentals are constant)

More Complex Situations: Trickle Down Theories (Anderson (1964))

The popularity of the "Rising Tide Raises All Boats" argument for the alleviation of poverty through growth (it is an irresistible temptation to comment that this is only true when all of the boats have no holes in them) has already been alluded to. For basic needs based definitions of the poverty cut-off this is no doubt true though it is not true if relative poverty is the measurement criterion. A slightly more sophisticated development of this argument is the "Trickle Down" effect (Anderson (1964)). The idea is that it is necessary for economic growth to initially benefit the higher income groups (because they make the marginal product of labour enhancing investments that increased the incomes of the poor) but it transits downward to the lower income groups over time. However this idea is predicated on the notion of economically different groups in society (with different investment behaviours for example) and it is not unreasonable to presume that different stochastic processes govern their respective behaviours. In effect one is "modeling" different groups, one of which would be the poor group, which will perhaps calls for a different approach to measuring poverty and inequality. The poor are identified by the extent to which their income processes are noticeably different from the income processes of other groups in society rather than because their income is less than some pre-specified boundary. It follows that some identifiably "rich" individuals may have, at least temporarily, incomes that are lower than some of the members in the poor group.

To explore the implications of this structure imagine the societal income process to be

that of a mixture of K normal distributions corresponding to the K income classes so that:

$$f_k(\ln(x_{kT})) = N((\ln(x_{k0}) + T(\mu_k + 0.5\sigma_k^2)), T\sigma_k^2) \quad k = 1,..,K$$

The classes are distinguished by their initial conditions in the sense that $\ln(x_{k0}) > \ln(x_{j0})$ for

all $j < k$ where $j, k = 1,..,K$ and the proportion of the population in each class is given by

w_k. Note that:

$$\int_0^x (f_i(z) - f_j(z))dz \leq 0 \ \forall \ x > 0 \ and \ i > j.$$

The class k=1 may be thought of as "the Permanently or Chronically Poor" so that CP,

the Chronic Poverty rate is w_1. In this form there will be chronically poor agents whose

incomes will be , at least temporarily, higher than some non-poor agents since the poor

and non-poor distributions will overlap. The extent to which the Permanently Poor

distribution overlaps the distribution of the other classes those members of those classes

may be considered the transitorily poor, so that the Transitorily Poor rate TP may be

written as:

$$TP = \sum_{k=2}^{K} w_k \int_0^\infty f_1(x_k)dx_k$$

For momentary convenience assume $\mu_k = \mu$ and $\sigma_k = \sigma$ for all k (i.e. the various classes are distinguished by their initial conditions alone). So here the poor are class k=1 and the rich are class k=K. Assume now that at period T_1 society moves to a new higher growth rate $\mu^* > \mu$, if all classes move together the income groups would have size distributions of the form:

$$\ln(x_{kT}) \sim N((\ln(x_{k0}) + T_1(\mu - \mu^*) + T(\mu^* + 0.5\sigma^2)), T\sigma^2)\ k = 1,.., K,\ and\ T > T_1$$

All classes retain their mean differences over time, the extent to which class A first order dominates class B remains constant and there is no change in the degree of polarization between the classes. However if in the first period only the highest income group (K) moves, j periods later only the next highest group (K-1) moves, j periods later only the next highest group (K-2) moves etc… Then the new societal income process will be the same mixture of K normal distributions corresponding to the K income classes but with:

$$\ln(x_{kT}) \sim N((\ln(x_{k0}) + (T_1 + jK - jk)(\mu - \mu^*) + T(\mu^* + 0.5\sigma^2)), T\sigma^2)\ k = 1,.., K,\ and\ T > T_1 + K$$

There are several observations to make. The distribution of average log incomes of the classes will be more widely spread initially and these initial differences will only be dissipated asymptotically (If the growth rates differ between the classes with the differences increasing with income class the differences will not dissipate asymptotically). In the short term the income size distribution of high income groups will more strongly first order dominate that of lower income groups increasing the polarization or lack of identification between poor and rich groups and the effect will be

larger the longer is the lag in the trickle down effect (j). In effect there will be greater absolute inequality in the short run. When poverty lines are defined relative to an income quantile this will increase the probabilities of both transient and chronic poverty for the lower income classes. All of this is predicated on all income groups differing only in their initial incomes, should there be heterogeneity in growth rates and variances as well as starting incomes then anything is possible.

It should be said that in this model structure so-called poverty cutoffs are superfluous since the poverty group is better defined as those agents governed by the process that is dominated by all other processes. Issues concerning measuring the plight of the poor would centre upon w_1, the mixture coefficient for the poor group, and measuring the differences between the poor group sub-distribution and the other distributions in the mixture. Inequality can also be measured in terms of these concepts or it can be measured in terms of the general variability characteristics (variance or coefficient of variation for example) of the overall mixture distribution. Similarly issues concerned with addressing the plight of the poor may then be seen in terms of influencing the weights attached to each class as well as changing the nature of the process that governs the poor group outcomes. The vector of weights in period t is given by w(t) and generational transitions take the form $w(t) = Tw(t-1)$ where $T - \|p(i|j)\|$ is a matrix of conditional probabilities of transiting to income group i given the agent was in group j where i, j = 1,.., K. $T (= JD^{-1})$ is a joint density matrix J post multiplied by an inverted diagonal matrix of marginal probabilities D^{-1}. In effect the policy makers have several instruments to work with, they can work with the parameters of the poor process (the various growth rates and

variability's in the poor processes), or they can work with T, getting agents to transit from poor processes to less poor processes in other words trying to move J towards an upper triangular matrix.

The Experience of Africa 1985-2005

To illustrate these issues data on per capita GDP for 47 African countries[9] together with

their populations were drawn from the World Bank African Development Indicators

CD-ROM for the years 1985, 1990, 1995, 2000, 2005 were used. An issue immediately

arises as to whether the raw data or population weighted data should be employed. At

the statistical level the parameters of interest will be estimated as though the data were

an independent random sample and the properties of those estimators predicated upon

that assumption. If the population of interest is that of Africa then this is not so and

sample weighting is necessary to adjust for the under sampling of highly populated

countries and over sampling of sparsely populated countries. A similar argument

prevails at the economic theoretic level if some sort of representative agent model is

presumed and the wellbeing of all Africans is of interest. For the purposes of

comparison, and to highlight the substantive differences the distinction makes both will

be reported here.

Table 1 reports summary statistics (means and variances) for the sample years and

clearly indicates the implicit over sampling of higher income nations in the un-weighted

---

[9] The countries in the sample were: Algeria, Angola, Benin, Botswana, Burkina Faso,
Burundi, Cameroon, Cape Verde, Central African Republic, Chad, Comoros, Congo,
Dem. Rep., Congo, Rep., Cote d'Ivoire, Egypt, Arab Rep., Equatorial Guinea, Ethiopia,
Gabon, Gambia, The, Ghana, Guinea, Guinea-Bissau, Kenya, Lesotho, Liberia,
Madagascar, Malawi, Mali, Mauritania, Mauritius, Morocco, Mozambique, Namibia,
Niger, Nigeria, Rwanda, Senegal, Seychelles, Sierra Leone, South Africa, Sudan,
Swaziland, Togo, Tunisia, Uganda, Zambia, Zimbabwe

estimates. Notice the un-weighted estimates record a growth of 14% over the period

whereas the weighted results report less than 8% growth, a substantial difference.

 Table 1. ln(GDP per capita)

	Un-weighted		Weighted	
Year	Means	Variances	Means	Variances
1985	6.1820771	0.92778351	6.0692785	0.90744347
1990	6.1784034	0.98733961	6.0827742	0.88023763
1995	6.1113916	1.1495402	6.0053685	0.95556079
2000	6.2174786	1.2386242	6.0488727	1.0276219
2005	6.3223021	1.3481553	6.1461927	1.0300315

Diagrams 2 and 3 present the beginning of period and end of period size distributions[10]

of ln(GDP per capita) again in un-weighted and weighted form. The shifts in location

and spread discerned in Table 1 can readily be perceived in these diagrams.

---
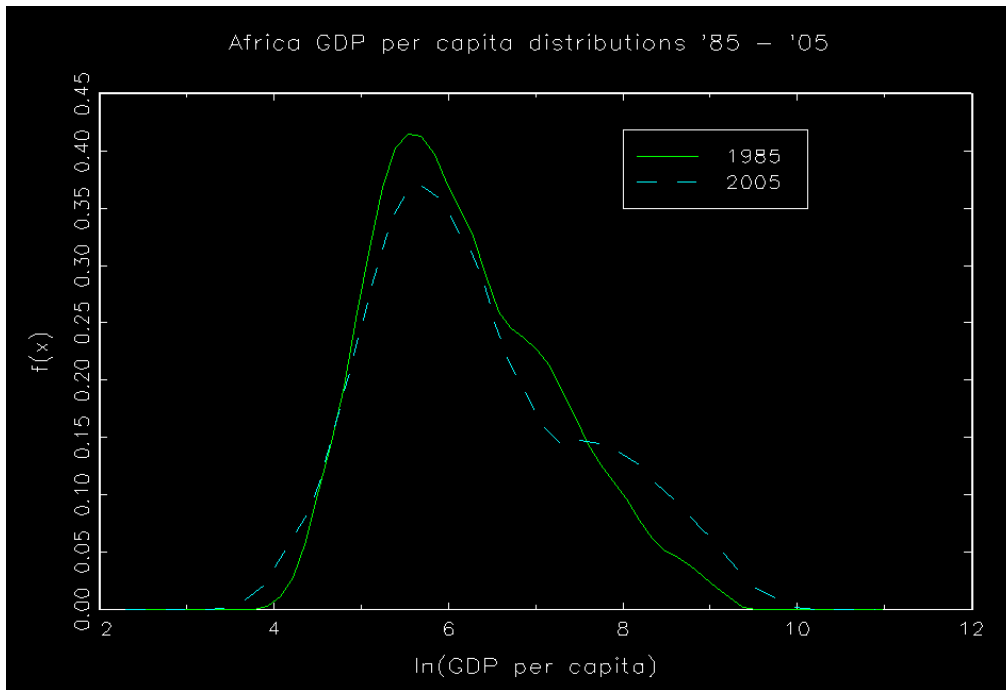
[10] These are essentially Epetchanikov kernel estimates of the respective size distributions.

Africa GDP per capita distributions '85 – '05 (population weighted)
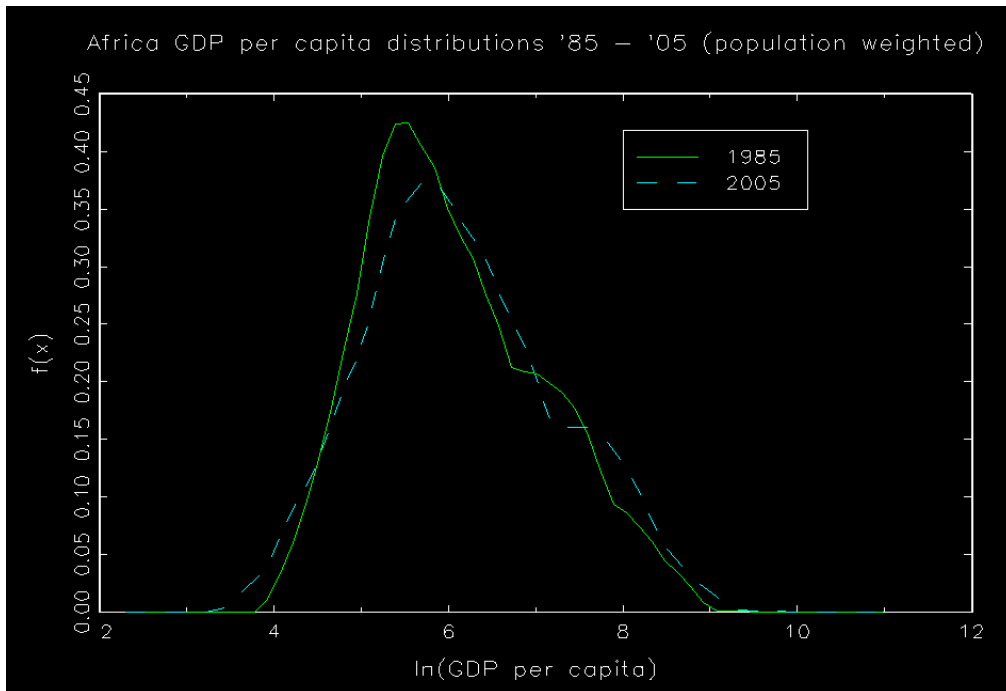
Pearson Goodness of Fit Tests of the hypothesis that these distributions are Log Normal
Or Pareto, performed for both weighted and unweighted samples, are reported in Table
2. At the 1% critical value there is a preponderance of evidence favouring the Log
Normal formulation (it only gets rejected twice in the un-weighted sample and once in
the weighted sample and pretty marginally so at that) whereas the Pareto gets solidly
rejected in every instance. This is slightly surprising since, to the eye, diagrams 1 and 2
suggest mixtures of 2 normals (one large poor group and a much smaller rich group),
but the evidence does not appear strong enough in the data to really reject pure
normality. Indeed the joint test of normality over the 5 observation periods does not
reject normality at the 1% level for the weighted sample. Interestingly enough the
populations are clearly log – normally distributed so that it may be inferred that gdp's
are themselves log- normally distributed since the difference or sum of two normally
distributed variables is also normally distributed.

Table 2.

Ln GNP per capita		
Un-weighted	Normal $\chi^2(4)$, [P(Upper Tail)]	Pareto $\chi^2(4)$, [P(Upper Tail)]
1985	12.320748 [0.015118869]	599.37137 [2.1196982e-128]
1990	18.172132 [0.0011420720]	648.71719 [4.4178901e-139]
1995	7.2879022 [0.12143386]	254.88561 [5.7677072e-054]
2000	10.035066 [0.039841126]	413.23135 [3.8477939e-088]
2005	19.439149 [0.00064419880]	418.01949 [3.5518925e-089]
All years $\chi^2(20)$	67.254997 [5.0783604e-007]	2334.2250 [0.00000000]
Weighted	Normal $\chi^2(4)$, [P(Upper Tail)]	Pareto $\chi^2(4)$, [P(Upper Tail)]
1985	5.4247526 [0.24642339]	850.67992 [8.0715931e-183]
1990	10.882507 [0.027916699]	890.92447 [1.5416785e-191]
1995	3.7072456 [0.44707297]	389.07173 [6.3881404e-083]
2000	2.8078575 [0.59047714]	869.74574 [5.9768551e-187]
2005	13.992590 [0.0073187433]	935.78289 [2.9403608e-201]
All years $\chi^2(20)$	36.814952 [0.012314284]	3936.2047 [0.00000000]
Ln Population		
Unweighted	Normal $\chi^2(4)$, [P(Upper Tail)]	Pareto $\chi^2(4)$, [P(Upper Tail)]
1985	2.6979621 0.60957127	194.39270 6.0291654e-041
1990	3.4815355 0.48069145	193.84403 7.9101293e-041
1995	2.5274974 0.63971883	308.86671 1.3243132e-065
2000	2.9924681 0.55908664	250.28037 5.6645137e-053
2005	3.7924813 0.43481828	245.77727 5.2865977e-052

Given the joint normality of the 5 observation periods is accepted, the restrictions

implied by [2] can be examined. The null and alternative hypotheses may be written as:

H_o: $f(x_{T+i}) \sim N((\ln(x_o)+(T+i)\mu+0.5\sigma^2),(T+i)\sigma^2)$ versus H_1: $f(x_{T+i}) \sim N(\mu_i,\sigma_i^2)$ for i=0, 1, …,4;

Table 3. Estimates and Tests of Restrictions

	Un-weighted Sample	Weighted Sample
$\ln(x_o)$	5.7827258	4.0511799
T (# of five year intervals)	9.9694699	103.01656
μ	-0.012055759	0.014659728
σ^2	0.094224026	0.0091376682
$X^2(6)$	0.19646781	0.78654360
$P(>\chi^2(6))$	0.99985320	0.99242963

Here $\ln(x_0)$ and T are parameters thus implying 6 restrictions on the alternative. Table 3

reports the estimates of $\ln(x_0)$, T, μ and σ^2 for the weighted and un-weighted samples,

together with the test of the restrictions. In both cases the restrictions are not rejected implying that, conditional on the underlying normality of the distributions, Gibrat's law is an adequate description of the data.

To examine what these different approaches imply for poverty measurement headcount measures of both the absolute and relative type are considered, the former based upon the popular dollar and 2 dollar a day cut offs (which for our sample become 5.8749307 Table 4.

Year	Year by year normality		Under Gibrats Law	
Absolute (1 and 2 dollar a day cut-off) Un-weighted Estimates				
1985	0.37491020	0.65569496	0.39532698	0.67353927
1990	0.38002613	0.65253180	0.38684395	0.65329716
1995	0.41272336	0.66492689	0.37893064	0.63472718
2000	0.37912189	0.62362764	0.37150161	0.61758498
2005	0.35000783	0.58381961	0.36449051	0.60167400
Absolute (1 and 2 dollar a day cut-off) Weighted Estimates				
1985	0.41916948	0.69972853	0.43567992	0.70969353
1990	0.41233910	0.69751462	0.42823308	0.70198016
1995	0.44692413	0.71757254	0.42088097	0.69425410
2000	0.43188059	0.69573737	0.41362387	0.68651911
2005	0.39462744	0.66118137	0.40646200	0.67877884
Relative (50% and 60% median cut-off) Un-weighted Estimates				
1985			0.23725235	0.29907775
1990			0.24768529	0.30767299
1995			0.25697797	0.31525458
2000			0.26532206	0.32200680
2005			0.27286739	0.32807031
Relative (50% and 60% median cut-off) Weighted Estimates				
1985			0.23748342	0.29926912
1990			0.23854875	0.30015083
1995			0.23960137	0.30102105
2000			0.24064154	0.30188005
2005			0.24166951	0.30272805

6.5680779 respectively) and the latter based upon the also popular 50% and 60% of the median cutoffs. As expected in general absolute poverty measures decline with economic growth and relative poverty measures decline.

As for mobility the transitions can be evaluated for each of the five year intervals via the distance of the joint density matrix from that of a diagonal. The statistic

$$\sum_{ij} \min(p_{ij}, \mathrm{diag}(p_{i.})).$$

provides an index of immobility, where p_{ij}, is the probability of a country being in category i in period t and in category j in period t+1, $\mathrm{diag}(p_{i.})$, a square matrix with the vector of probabilities of being in category i in period t on the on diagonal corresponds to complete immobility (Anderson and Leo (2008) establish that this is asymptotically normal). Splitting this sample into 5 equal sized categories the four 5 year transitions generate statistics (standard errors) 0.87234 (0.04868) for the first thre transitions and 0.82979 (0.05482) for the fourth. This corresponds to very high immobility consistent with Gibrats law though it was expected to be increasing rather than staying constant.

The African Distribution as a mixture of normals.

As observed earlier the diagrams are very suggestive of a mixture of normals one largish poor group and one smaller rich group and it is of interest to see the consequences of modelling the processes under this structure. First it is appropriate to examine the degree of mobility within the distribution over time. Table 5 reports the 20 year transition probability matrix and indicates that in essence there appears to be very little mobility over the period between 5 rank groups (the five rank cells were 1-10 11-19 20-28 29-37 38-47). Essentially 5 countries moved up from cell 1 to cell 2 and 4 moved down from cell 2 to cell 1, 1 moved up from cell 2 to cell 3, two moved up from cell 3 to cell 4

and one moved down from cell 4 to cell 3. The only change of more than one cell was

Liberia who dropped from cell 4 to cell 1 over the period and one other original cell 4

Table 5. 20 year Transition Probability Matrix

1985 Cell	2005 Cell				
	1-10	11-19	20-28	29-37	38-47
1-10	0.10638298	0.10638298	0.00000000	0.00000000	0.00000000
11-19	0.085106383	0.085106383	0.021276596	0.00000000	0.00000000
20-28	0.00000000	0.00000000	0.14893617	0.042553191	0.00000000
29-37	0.021276596	0.00000000	0.021276596	0.12765957	0.021276596
38-47	0.00000000	0.00000000	0.00000000	0.021276596	0.19148936

Immobility Index   0.65957447   Standard Error 0.069118460

Member moved up to cell 5 and one cell 5 member moved down to cell 4. In sum there

appears to be some deal of mobility at the lowest end of the spectrum but very little

elsewhere, certainly it is reasonable to assume that memberships of the large poor and

small rich groups apparent in diagrams 1 and 2 (and hence the mixture coefficients)

appear to be relatively constant.

Techniques for estimating mixtures of normals are available (see for example Johnson

Kotz and Balakrishnan (1994)) but tend to be complex and depend upon fairly large

numbers of observations. Here, since there are a limited number of observations, an ad

hoc method is used for simplicity and convenience, but it turns out to be quite successful

in terms of replicating the empirical distribution . Given the evidence is that the

membership of the groups is very stable over the period, countries are allocated into rich

and poor groups in the following fashion. Visual inspection of the 2005 distribution in

diagram 1 suggests that the modal values of the poor and rich groups are approximately

6 and 7 respectively. Observations below 6 can be almost all be attributed to the poor

group and similarly observations above 7 can be similarly attributed to the rich group

and the corresponding observations were allocated accordingly. Given the symmetry of

the underlying log - normals around their respective modes, the areas under the curve

corresponding to these two regions reflects the relative size and hence weights w_r

Disposition of Poor and Rich Countries.

Poor Group	Rich Group
Benin, Burkina Faso, Burundi, Cape Verde, Central African Republic, Chad, Democratic republic of the Congo, Cote d'Ivoire, Ethiopia, The Gambia, Ghana, Guinea-Bissau, Kenya, Liberia, Madagascar, Malawi, Mali, Mauritania, Mozambique, Niger, Rwanda, Sierra Leone, Sudan, Togo, Uganda, Zambia, Zimbabwe.	Algeria, Angola, Botswana, Cameroon, Comoros, Republic of the Congo, Egypt, Equatorial Guinea, Guinea, Lesotho, Mauritius, Morocco, Namibia, Nigeria, Senegal, Seychelles, South Africa, Swaziland, Tunisia.

(20/47) and w_P (27/47) of the rich and poor groups. The observations between 6 and 7

were allocated randomly according to these weights to the rich and poor groups. After

an initial fit of the individual poor and rich country distributions a below median poor

country was switched with an above median rich country[11] which improved the fit so

that the following two rich and poor subgroups were established.

Having partitioned the sample in this fashion estimation of the mixture distribution is

quite simple in both unweighted and population weighted modes. Tables 6 and 7 and

diagrams 3 and 4 report the results. In both cases the fits are extremely good and

correspond to a more than adequate description of the data. The poor group has enjoyed

zero economic growth and the rich group has enjoyed a steady one percent annual

growth rate over the period. Differences between the un-weighted and weighted cases

emerge when gdp per capita levels and variabilities are concerned. In the unweighted

---

[11] Relative to a normal distribution the initial poor country distribution appeared attenuated in the upper tail
and the rich country distribution appeared attenuated in the lower tail.

case income levels are generally higher and variances are lower but increasing over time whereas in the weighted case incomes are lower and variances are higher but diminishing over time. The restrictions implied by Gibrat's law for the separate poor and rich groups in both weighted and unweighted samples are rejected in all cases (frequently resulting in nonsense estimates such as negative variances and negative time parameters) though basic log normality is not rejected in any case suggesting that Kalecki's Law is the best description of the data for the individual groups.

Table 6. A Mixture of 2 Log Normals (Poor group and Rich group).

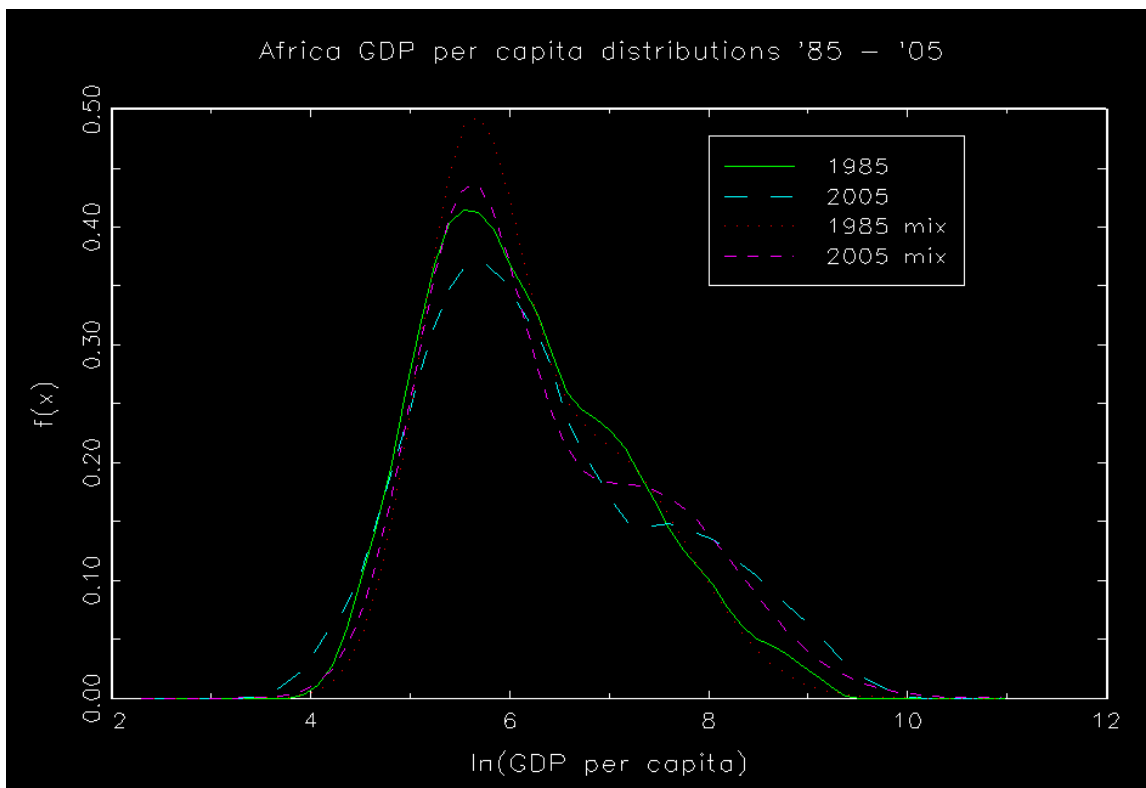|	Poor Mean	Rich Mean	Poor Variance	Rich Variance	$\chi^2(4)$	[P(Upper Tail)]
1985	5.5869	6.9856	0.2691	0.6950	3.6312	0.4582
1990	5.5425	7.0369	0.2350	0.7183	5.9701	0.2014
1995	5.4325	7.0278	0.3461	0.7705	0.2800	0.9911
2000	5.5006	7.1852	0.3258	0.8370	0.4526	0.9780
2005	5.5680	7.3406	0.3219	0.9234	1.9169	0.7510

Table 7. Sample Weighted Mixtures

	Poor Mean	Rich Mean	Poor Variance	Rich Variance	$X^2(4)$	[P(Upper Tail)]
1985	5.4115	6.7584	0.6262	1.6988	3.0473	0.5499
1990	5.4146	6.7975	0.5422	1.4666	1.6053	0.8078
1995	5.3109	6.7532	0.5743	1.4099	4.3005	0.3669
2000	5.3423	6.8289	0.5671	1.2868	8.2530	0.0827
2005	5.4340	6.9527	0.4943	1.1327	5.9860	0.2002
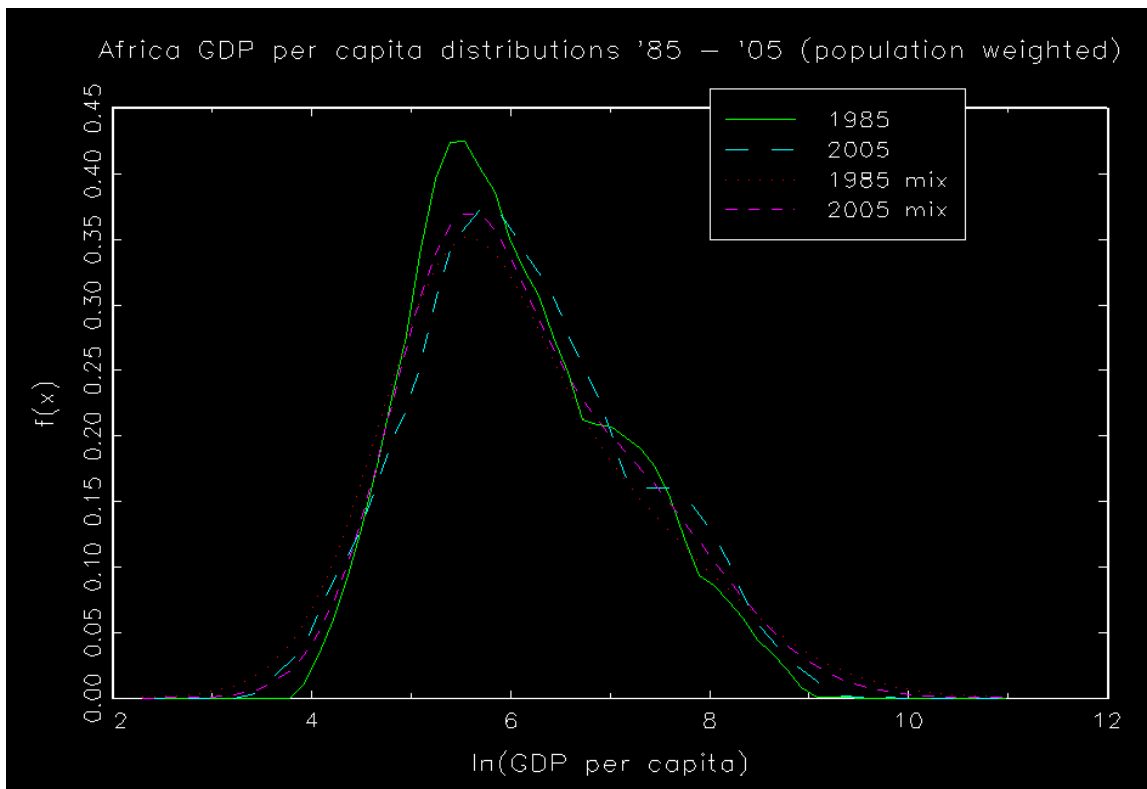
Poor and rich groups appear to be moving apart, Table 8 reports a trapezoidal measure of bi-polarization (Anderson (2008)) for both weighted and un-weighted samples illustrating the point. This may be interpreted as the poor becoming relatively poorer

Table 8. Bi-Polarization Index = $0.5(f_p(x_{pmode}) + f_r(x_{rmode})) (x_{rmode} - x_{pmode})$

Year	Unweighted Sample		Weighted Sample	
1985	0.87249993	(0.027374440)	0.54564641	(0.046480295)
1990	0.96662956	(0.028177909)	0.60239981	(0.042441864)
1995	0.90342925	(0.028177909)	0.62192838	(0.041522365)

2000	0.95600414	(0.030424419)	0.65517946	(0.039377304)
2005	0.99116013	(0.032321510)	0.71551955	(0.036396440)

In this circumstance the issue of population weighting makes a big difference, with no

population weighting the poor group are becoming absolutely poorer and exhibiting

diminishing within group association, the source of polarization is the increased

alienation or distance between the two groups. With population weighting they are not

Table 9. Polarization Tests[*]

Unweighted Sample

Comparison Years	Difference	("t"test)	{P(T<t)}
1990-1985	0.094129622	(2.3960367)	{0.99171328}
1995-1985	0.030929320	(0.77595362)	{0.78111181}
1995-1990	-0.063200302	(-1.5637535)	{0.058937739}
2000-1985	0.083504205	(2.0403293)	{0.97934123}
2000-1990	-0.010625417	(-0.25622838)	{0.39888725}
2000-1995	0.052574885	(1.2513981)	{0.89460536}
2005-1985	0.11866019	(2.8014895)	{0.99745663}
2005-1990	0.024530571	(0.57207798)	{0.71636543}
2005-1995	0.087730873	(2.0211482)	{0.97836779}
2005-2000	0.035155988	(0.79200942)	{0.78582241}

Weighted Sample

Comparison Years	Difference	("t"test)	{P(T<t)}
1990-1985	0.056753398	(0.90167380)	{0.81638491}
1995-1985	0.076281972	(1.2239198)	{0.88950876}
1995-1990	0.019528574	(0.32890099)	{0.62888474}
2000-1985	0.10953305	(1.7980440)	{0.96391498}
2000-1990	0.052779657	(0.91163730)	{0.81902016}
2000-1995	0.033251083	(0.58106062)	{0.71940020}
2005-1985	0.16987315	(2.8775040)	{0.99799583}
2005-1990	0.11311975	(2.0232198)	{0.97847475}
2005-1995	0.093591173	(1.6950007)	{0.95496236}
2005-2000	0.060340090	(1.1252951)	{0.86976799}

* Tests are based on the trapezoid measure being asymptotically normally distributed
with a variance $\approx (f(x_{1m})+f(x_{2m}))^2(f(x_{1m})/[f''(x_{1m})]^2+f(x_{2m})/[f''(x_{2m})]^2)||K'||^2_2$ where x_{mj} j =
1,2 are the modes of the respective distributions, where f() is the normal and K is the
Gaussian kernel (Anderson, Linton and Wang (2008)).

becoming absolutely poorer but are exhibiting increased within group association, there is a small amount of between group alienation but a substantial increase in within group association. The only significant changes in polarization in both comparison types were increases in polarization over time. In both cases the poor and rich groups are following distinct stochastic processes and there is no sense in which "the rising tide is raising all boats" or improvements in the well being of the rich African countries are trickling down to the poor countries.

Conclusions.

It is not at all clear that boats and tides aphorisms and trickle down theories apply either in theory or practice when the well-being indicator is well described by some sort of stochastic process, especially when the process is one that is frequently observed in practice. It really depends on the nature of poverty or inequality being considered as well as the precise nature of the stochastic process(es) involved. Stochastic processes that are non-stationary engender distributions whose dispersion (absolute inequality) increases over time, whether or not relative inequality increases depends upon the nature of the growth process. Similar statements can be made about poverty, but here the nature of the growth process affects both absolute and relative poverty. What is clear is that it is not unequivocally the case that rising tides raise all boats or that wellbeing unequivocally trickles down even in the simplest of circumstances. This is even more so the case when the progress of the poor and the non-poor are described by different stochastic processes.

In the case of Africa when GNP per capita is modelled over the recent two decades as a singular stochastic process the prediction of Gibrat's law appears to hold true regardless of whether the analysis is performed under a population weighting scheme or a non-weighted scheme in the sense that the distribution is log normal. Under this description absolute poverty is diminishing and relative poverty is increasing and absolute inequality is increasing and relative inequality is diminishing. Kernel estimates of the density indicated some evidence of bimodality suggesting a mixture of at least two distributions. When log GNP per capita is described by a mixture of two normals (which was not rejected by the data), one describing the poor country process and the other describing the rich country process, it is apparent that the two groups are polarizing, with the poor group in this sense becoming relatively poorer. In this circumstance the issue of population weighting made a big difference, with no population weighting the poor group are becoming absolutely poorer and exhibiting diminishing within group association, with population weighting they are not becoming absolutely poorer but are exhibiting increased within group association.

References.

Anderson, W.H.L. (1964) "Trickling Down, The relationship between Economic Growth and the Extent of Poverty of American families" Quarterly Journal of Economics 78, 511-524.

Anderson G.J. , Linton O. and Wang Y. (2009) "Non-Parametric Estimation of Polarization Measures" Mimeo LSE.

Anderson G.J. (2008) "Polarization of the Poor: Multivariate Relative Poverty Measurement Sans Frontiers" Mimeo University of Toronto Economics Department

Auerbach, F., 1913, Das Gesetz der Belvolkerungskoncertration, Petermanns Geographische Mitteilungen 59, 74-76.

Champernowne D.G. (1953) A Model of Income Distribution Economic Journal 63 318-351.

Deaton A. and C. Paxson (1994) Ïntertemporal Choice and Inequality"Journal of Political Economy 102 384-394.

Duclos J-Y., J. Esteban and D. Ray (2004) "Polarization: Concepts, Measurement and Estimation" Econometrica 72, 2004, 1737-1772.

Durlauf, S., P. Johnson and J. Temple (2005) "Growth Econometrics"forthcoming in Handbook of Growth Economics. P. Aghion and S Durlauf editors.

Ferguson A. (1767) A History of Civil Society.

Gabaix, X., 1999, Zipf's law for cities: an explanation, Quarterly Journal of Economics 114, 739-767.

Gibrat, R., 1930, Une Loi Des Repartitions Economiques: L'effet Proportionelle, Bulletin de Statistique General, France, 19, 469.

Gibrat, R., 1931, Les Inegalites Economiques (Libraire du Recueil Sirey, Paris).

Gnedenko B.V. (1962) Theory of Probability Chelsea New York.

Harrison A. (1984) "A Tale of Two Distributions" Review of Economic Studies

Harrison T., (1985) "Brownian Motion and StochasticFlow Systems" Malabar F.C. Kreiger.

Johnson, Kotz and Balakrishnan (1994).

Kalecki M., (1945), On the Gibrat distribution, Econometrica 13, 161- 170.

Meghir C and C Pistaferri (2004) Ïncome Variance Dynamics and Heterogeneity"
Econometrica 72 1-32.

Meyer B and J. Sullivan (2002) "Measuring the Wellbeing of the Poor Using Income and
Consumption" Journal of Human Resources 1180-1220.

Neal D. and S. Rosen (2000) "Theories of the Distribution of Earnings" Chapter 8 in
Handbook of Income Distribution Volume 1 A.B. Atkinson and F. Bourguignon eds.
North Holland

O'Neill D., (2005) "The Welfare Implications of Growth Regressions"Mimeo.
Pareto, V., 1897, Cours d'Economie Politique (Rouge et Cie, Paris).

Pearson, K., 1900, On a criterion that a given system of deviations from the probable in
the case of a correlated system of variables is such that it can reasonably be supposed to
have arisen from random sampling, Philosophical Magazine 50, 157-175.

Reed, W. J., 2001, The Pareto, Zipf and other power laws, Economics Letters 74, 15-19.

Sen A.K. (1983) "Poor Relatively Speaking" Oxford Economic Papers 35 153-169

Silverman B.W. (1986) Density estimation for Statistics and Data Analysis. Chapman
Hall.

Simon, H., 1955, On a class of skew distribution functions, Biometrica 42, 425-440.

Smith A. (1776) An Enquiry Into the Nature and Causes of the Wealth of Nations.
Liberty Classics

Steindl, J., 1965, Random processes and the growth of firms (Hafner, New York).
Whittle, P. 1970, Probability (Wiley).

Townsend (1985) "A Sociological Approach to the Measurement of Poverty – A
Rejoinder to Professor Amartya Sen" Oxford Economic Papers 37 659-668.

Zipf, G., 1949, Human behavior and the principle of last effort (Cambridge MA: Addison

Wesley)

Appendix 1: Some variations on this theme.

1) Allowing T to be a random variable described by a geometric distribution of the form:

$$P(T = t) = p(1 - p)^{t-1}; t = 1, 2, ...$$

Reed (2006) shows that x will have a pdf f(x) which, if $P(\delta > 0) > 0$, has an upper tail such that:

$$f(x) \sim c_1 x^{-\alpha-1} \quad with \quad 1 - F(x) \sim c_2 x^{-\alpha}$$

In addition if $P(\delta < 0) > 0$, he shows that f(x) has a lower tail such that:

$$f(x) \sim c_3 x^{\beta-1} \quad with \quad F(x) \sim c_4 x^{\beta}$$

All with c_1, c_2, c_3, c_4, α and β positive. In effect f(x) is a Double Pareto distribution with some interesting properties. The distribution has a tent-like appearance and the thickness of the tails is essentially governed by the values of α and β in that when they are small the tails will be longer and fatter. The lower tail parameter (α) will be smaller when (p/(1-p)) is small and when individual income growth (δ) is larger with high volatility (σ). The upper tail parameter (β) will be small when p/(1-p) is small and income growth is low with high volatility. Thus high p/(1-p) and low income growth volatility causes both parameters to be large. A high average income growth rate implies a long upper tail and a short lower tail.

2) Allowing the initial size to be a log normal variate such that

$$\ln(x_0(t)) \sim N(a + t(\mu_0 - 0.5\sigma_0^2), T\sigma_0^2)$$

It can be shown that income of an agent of age T at time τ will be log normally distributed with:

$$E(\log(X)) = A_0 + (\tilde{\mu} - 0.5\tilde{\sigma}^2)T$$
$$V(\log(X)) = B_0^2 + \tilde{\sigma}^2 T$$

where A_0 and B_0^2 are the mean and variance of the starting size. Allowing time T from the initiation of the process to be exponential Reed (2003) showed the distribution of income sizes at a point in time to be the product of independent and double Pareto and log normal components with f(x) of the form:

$$f(x) = \frac{\alpha\beta}{\alpha+\beta}\left[x^{-\alpha-1}e^{\alpha v + \alpha^2\tau^2/2}\Phi\left(\frac{\log x - v - \alpha v^2}{\tau}\right) + x^{\beta-1}e^{-\beta v + \beta^2\tau^2/2}(1 - \Phi\left(\frac{\log x - v - \beta v^2}{\tau}\right))\right]$$

which he names the double Pareto log normal (dPln) distribution, it has a cdf of the form:

$$F(x) = \Phi\left(\frac{\log x - v}{\tau}\right)$$
$$-\frac{1}{\alpha+\beta}\left[\beta x^{-\alpha}e^{\alpha v + \alpha^2\tau^2/2}\Phi\left(\frac{\log x - v - \alpha v^2}{\tau}\right) + \alpha x^{\beta}e^{-\beta v + \beta^2\tau^2/2}(1 - \Phi\left(\frac{\log x - v - \beta v^2}{\tau}\right))\right]$$

where α, β, υ and τ are parameters of interest. As α and β tend to ∞ the distribution tends to log normal and as τ tends to 0 the distribution tends to double Pareto.