“Convergence” Hypotheses are Ill-Posed:
Non-stationarity of Cross-Country Income Distribution Dynamics

Mahmoud A. El-Gamal and Deockhyun Ryu*

January 2004

Abstract

The recent literature on “convergence” of cross-country per capita incomes has been dominated by two competing hypotheses: “global convergence” and “club-convergence”. This debate has recently relied on the study of limiting distributions of estimated income distribution dynamics. Utilizing new measures of “stochastic stability”, we establish two stylized facts that question the fruitfulness of the literature’s focus on asymptotic income distributions. The first stylized fact is non-stationarity of transition dynamics, in the sense of changing transition kernels, which renders all “convergence” hypotheses that make long-term predictions on income distribution, based on relatively short time series, less meaningful. The second stylized fact is the periodic emergence, disappearance, and re-emergence of a “stochastically stable” middle-income group. We show that the probability of escaping a low-income poverty-trap depends on the existence of such a stable middle income group. While this does not answer the perennial questions about long-term effects of globalization on the cross-country income distribution, it does shed some light on the types of environments that are conducive to narrowing/widening the gap between rich and poor countries.

Keywords: global income distribution; convergence clubs; transition kernel; stochastic stability.

* El-Gamal is a Professor of Economics and Statistics at Rice University, where Ryu is a Ph.D candidate in the Economics Department. Contact: Mahmoud A. El-Gamal, Economics MS 22, Rice University, 6100 Main Street, Houston, TX 77005. Emails: elgamal@rice.edu, dhryu@rice.edu.

1. Introduction

In the age of globalization, no issue is more hotly debated than the issue of the income gap between rich and poor countries. The neoclassical growth model (with well-behaved aggregate production functions) suggests that the gap between rich and poor countries should be collapsing over time. This was rephrased in the early empirical literature in terms of two types of "convergence hypotheses", named absolute and conditional convergence. Early empirical results supported those convergence hypotheses, by estimating negative coefficients for "initial income" variables in income growth regressions, c.f. Mankiw et al (1992), Barro and Sala-i-Martin (1992, 1995).

The simple convergence story was soon challenged both on the theoretical and empirical levels. At the theoretical level, endogenous growth models challenged the credibility of – and robustness to – neoclassical assumptions of well-behaved aggregate production functions. With minor modifications of the Solow aggregate production function, non-convexities and poverty traps become possible, as shown in threshold models of Azariadis and Drazen (1990), Durlauf (1993), and others. Meanwhile, at the empirical level, the results of simple cross-country growth regressions that lent support to convergence hypotheses were challenged by the results of panel data, time series (unit root and co-integration tests) and distributional dynamics analyses.[1]

Most recently, the distributional dynamics approach has become increasingly popular, e.g. see Quah (1993a, 1993b, 1995, 1996a, 1996b, 1997, 2001) as well as Bianchi (1997), Desdoigts (1999), Johnson (2001), Bulli (2001). Unlike the simple linear regression approach, the distributional dynamics literature focuses on richer (non-linear) dynamics as captured by a transition probability matrix for world income distribution groups. The findings of this literature reject the simple convergence hypothesis, and point to multi-modality in the limiting distribution of per capita incomes. Durlauf (1993) also found evidence of this so-called "club convergence" by introducing non-linearity through a regression-tree method, which allowed for the possibility of different countries following different growth dynamics.

Despite utilizing different methodologies, and reaching very different conclusions, the linear regression convergence literature and the distribution dynamics club convergence literature share a common weakness. Both approaches purport to make conclusions regarding long-term dynamics, while utilizing relatively short time series. This asymptotic approach is particularly problematic in the linear regression literature, which assumes that all countries share the same linear growth dynamic, in essence forcing the estimated parameters to support the convergence hypothesis[2] (since explosive

[1] For a survey of early developments in the empirical growth literature, see Durlauf and Quah (1999).

[2] In this regard, Phillips and Sul (2003) show that obtaining β-convergence and/or α-convergence is not possible if economies have different speeds of convergence, different growth rates due to technological progress, and heterogeneous initial income.

divergence and cyclicality – the only other dynamics allowed by linear models – are very far fetched). The club convergence literature is also problematic, since it focuses on the existence of multi-modal unique invariant measures to which income distributions must converge. However, this invariant measure is consistent with the U.S. always being significantly richer than Bangladesh, as well as the scenario wherein the two countries trade places every period. Moreover, as we shall see in this paper, there is evidence of fundamental non-stationarity of transition dynamics, which make any asymptotic predictions based on short time-series difficult to support.

In this paper, we focus on the estimated income dynamics themselves, rather than their supposed limiting distributions. This approach allows us to capture rich nonlinearities in income dynamics, while avoiding the anonymity and assumed stationarity of stochastic dynamics that are imposed by focusing on ergodic distributions of estimated transition kernels. Our approach is based on the following notion of stochastic stability: the relative likelihood for any given country at its particular relative income of getting relatively richer, relatively poorer, or staying the same relative income. A stochastically stable point on the income distribution scale is a ranking with a "stochastic basin of attraction", whereby countries that are slightly richer are likely (probability > 0.5) to get poorer, and those slightly poorer are likely to get richer. Our analysis of "clubs" is therefore reduced to a statistical analysis of the number of such stochastically stable points, one for which we derive a formal statistical test.

In the distributional dynamics literature, the initial dominant result, studying ergodic distributions (unique invariant measures) of estimated income distribution transition matrices, has been bi-modality in the limiting income distribution.[3] If a middle group was in fact observed in any historical time period, this asymptotic result was explained in terms of a "vanishing middle" in the income distribution, in the sense that middle income countries will have eventually to fall to either extreme (joining the rich club or the poor club, or – theoretically – oscillating between the two!). More recent studies of distributional dynamics, e.g. Kremer, Onatski, and Stock (2001, hereafter: KOS), found evidence of uni-modality of the limiting distribution.[4] Sala-i-Martin (2002a, 2002b) also found evidence for asymptotic uni-modality (a variant on the classical convergence hypothesis) by utilizing a controversial data transformation.[5]

[3] Quah (1997) describes this result as evidence for a polarization of the world income distribution and vanishing of middle-income group. Bianchi (1997), Paap and van Dijk (1998) also support the conclusion of twin-peaks and vanishing middle-income group.

[4] Quah (2001) provided the counter-argument that KOS (2001)'s single-peaked limit distribution would be reached only after centuries of polarization. However, that critique merely points to the general weakness of any analysis that focuses primarily on limiting distributions (including Quah's own studies), rather than studying the actual estimated short to medium-term dynamics.

[5] Bianchi (1997) correctly points out that Sala-i-Martin's logarithmic transformation simplifies the structure of income distribution data by removing skewness and outliers, thus making it more difficult to distinguish between a bimodal density and a unimodal one.

Quah (2001) pointed out that the literature on distributional dynamics had to-date relied mainly on ocular inspection of estimated ergodic measures, and lacked a formal test of the club convergence hypothesis. The best attempts to-date at providing a formal test of multi-modality are those of Bianchi (1997), and KOS (2001). However, Bianchi (1997) merely tested for multi-modality of the static income distribution density. In other words, this test only applies to the case where income distribution is already drawn from the unique invariant (ergodic) measure. In Appendix B, we provide a dynamic version of this test, inspired by the methodology used by KOS (2001). The latter test, by the admission of its own authors, suffers from two fundamental weaknesses: non-robustness to the supposed locations of the "clubs", and ability only to test for one club versus more. The stylized facts uncovered in Section 2 question the validity of focusing on limiting distributions, in light of the fundamental non-stationarity of income distribution dynamics.

This apparent non-stationarity of income distribution dynamics (in the stochastic processes sense of changing transition probabilities/kernels) prompted us to abandon the focus on limiting distributions of estimated transition kernels (although we perform some limiting distribution analyses, for comparison purposes). Instead, we focus on the transition dynamics in different sub-periods, to investigate our notion of stochastic stability of various income groups. Utilizing the seminal works on time-series nonparametric density estimation by Roussas (1969a, 1969b, 1991), we develop a formal test of multiple stochastically stable income groups in different sub-periods

The rest of this paper will proceed as follows. In Section 2, we establish empirical stylized facts regarding the periodic appearance, disappearance and re-appearance of a middle-income group, as well as non-stationarity of income distribution dynamics. In section 3, we propose a notion of stochastic stability for studying the properties of transition dynamics during those sub-periods within which a middle group disappeared or re-emerged. We also derive the asymptotic distribution theory for our measure of stochastic stability, thus paving the road for a formal statistical test of the stability of multiple income-distribution clubs, which is presented in Section 4. We conclude the paper in Section 5 with discussions of the economic implications of our stylized facts and empirical results.

2. Empirical Stylized Facts

We would like to provide graphical evidence of two main stylized facts in income distribution dynamics. The first is periodic appearance, disappearance and re-appearance of middle-income groups, and the second is the periodic change (non-stationarity) in the transition kernel itself. The latter stylized fact puts into question the study of "convergence", "club convergence", or any hypothesis based on the assumption of constant transition dynamics. In other words, the central question of the convergence literature may be ill-posed in light of this apparent, and fundamental, non-stationarity.

The data we utilize in this paper is taken from the Summers and Heston Penn World Table Mark 6.1.[6] Our measure of a country's position in the global income distribution is the standard relative income, calculated using PPP-based per capita incomes of the various countries.[7]

We must mention at this juncture that a number of recent studies have suggested replacing national per capita GDPs, as the object of investigation of income distribution dynamics, with global individual incomes (cf. Bourguignon and Morrison (2002), Fischer (2002), and Sala-i-Martin (2002a, 2002b)). With the exception of Bourguignon and Morrison (2002)'s results, this alternative object of investigation provides support for the classical convergence hypothesis, mainly by giving China and India (accounting for one-third of the world-population, and experiencing some of the fastest rates of per capita income growth in the world) more weight. Proponents of this alternative object of investigation criticized per capita GDP as an oversimplified measure that ignores within-country inequality effect, c.f. Bourguignon and Morrison (2002). Thus, Fischer (2002) finds evidence for classical β-convergence (negative estimated coefficient for initial per capita income in a growth regression) by using population-weighted real per capita GDP as the variable of interest. Sala-i-Martin (2002a and 2002b) used direct kernel estimates of global income distribution to account for within–country reductions of income distribution inequality. He concluded that world income inequality was on the decline, thus supporting the hypotheses of "vanishing twin peaks" and "emergence of a world middle-class", c.f. Sala-i-Martin (2002b), p.14.

This new empirical literature notwithstanding, we have decided, for a number of reasons, to limit our analysis to the dynamics of cross-country per capita GDP distributions: First, the "convergence" literature was originally motivated by the predictions of the neo-Classical Solow growth model, based on its well-behaved aggregate production function. Second, the bulk of the literature on income distribution dynamics to-date focused on distributions of cross country per capita GDP. This paper primarily criticizes that literature by establishing stylized facts regarding the fundamental non-stationarity of the transition dynamics of those distributions. This fundamental non-stationarity makes claims regarding all "convergence" hypotheses of doubtful relevance. Third, we would argue that future research on within- as well as between-country income distributions should consider the possible fundamental non-stationarity of within-country income distribution dynamics as well. Finally, the recent literature, e.g. Bourguignon and Morrison (2002, the Table 2) and Sala-i-Martin (2002a. p.39), suggests that between-country income inequality effects dominate within-country effects. Thus, we do not distort the results significantly by focusing on the former.

---

[6] Available at http://pwt.econ.upenn.edu.

[7] As commonly done in this literature, we define relative income as a percentage share of total world income. Relative income x_i is thus calculated as follows: $x_i = (X_i / \sum X_i) \times 100$, where X_i is per capita GDP of a country.

To establish our first stylized fact, we employ two different graphical representations of the data. First, we plot time series of cross-country per capita GDP Gini coefficients.[8] Classical convergence hypotheses predict that Gini coefficient would be decreasing over the time, while divergence hypotheses predict that they would be trending upwards. The first graph in Figure 1 shows that the Gini coefficients for our entire dataset (including developed and developing countries) trend upwards. On the other hand, the third graph in the left panel in Figure 1 shows that the Gini coefficient within the group of OECD countries is declining over the time. Those two graphs support the two-clubs convergence hypothesis. Further dividing the non-OECD group into three geographical sub-groups (graphs on the right panel in Figure 1): African, Latin American, and East and South-east Asian, we find evidence of non-monotonicity. Those sub-groups include Latin American and Asian countries that occupied a middle-income position for various sub-samples. The non-monotonicity of their Gini coefficients presents a first hint at the emergence, disappearance, and re-appearance of middle income groups.[9]

To investigate the emergence and stability of this middle income group, we propose a measure of "stochastic stability" for a country at some relative income level. Our measure is constructed as the median of the given country's relative income at time *t*+1 conditional on its relative income at time *t*, less that country's relative income at time *t*. In other words, denoting the period *t* relative income of a given country by x_t, we consider the measure:

$$f(a) = \text{median}(x_{t+1} \mid x_t = a) - a.$$

A zero of this function occurs at any value *a* such that a country with that relative income is equally likely to move up in the income distribution as it is to move down. Plots of this measure are qualitatively similar to those used by Quah (1997) and Johnson (2000) to investigate income distribution dynamics graphically. However, as we shall see shortly, a distribution theory is relatively easy to derive for our proposed measure.

We now proceed to define the "stochastic stability" of a zero of our function $f(a)$. Towards that end, it is clear that whenever $f(b)$ is positive, a country with relative income *b* is likely (probability > 0.5) to move up in the income distribution, and *vice versa*. Consequently, a stochastically stable zero of $f(.)$ is defined as a point at which the function crosses the x-axis from above. For instance, Figure 2 illustrates a hypothetical function $f(.)$ with five zeros. Three of the function's crossings are from

---

[8] Computed, as usual as: $G = \dfrac{\sum_{i=1}^{n}\sum_{j=1}^{n}\left|x_i - x_j\right|}{2n^2\mu}$,

where μ is mean size, c.f. http://mathworld.wolfram.com/GiniCoefficient.html.

[9] Further sub-dividing African countries into four regional sub-groups (western, central, eastern, southern), we detect more within-group heterogeneity of Gini coefficients. Sala-i-Martin (2002a) has clearly made the case for the importance of understanding the various income dynamics within Africa, as a tool for better understanding global income dynamics.

above (C1, C3, and C5), i.e. “stochastically stable”, in the sense that countries slightly to the left of those points are likely to move to the right, and *vice versa*. In contrast, C2 and C4, wherein the function crosses the x-axis from below, are “stochastically unstable”. Focusing on short-term stochastic stability, rather than asymptotic distributions, we may define an analogue of Quah’s “twin peaks” hypothesis in terms of a function $f(.)$ with two stochastically stable zeros and a “vanishing middle”. In contrast, a function with three stochastically stable zeros would indicate that there are three clubs, the middle one being Sal-i-Martin’s “emerging middle class”. Notice in this regard that the function $f(.)$ is derived directly from the transition kernel, and not merely from its hypothesized asymptotic limit.

Utilizing plots of our proposed function, we can establish both stylized facts mentioned at the beginning of this section. In Figure 3, we show plots of nonparametric estimates of the function $f(.)$ for the two sub-periods: 1961-66 and 1991-96. We shall describe our nonparametric method for estimating $f(.)$, and the asymptotic statistical properties of the resulting estimate, in Section 3. Without constructing confidence intervals around our estimates, however, it seems compelling to conclude that there were two stable zeros in the early 1960s, but three stable zeros in the early 1990s. Not surprisingly, the so-called “Asian Tigers” occupy the middle income group in the early 1990s. Indeed, as we shall see in the graphs accompanying the formal tests in Tables 4-10, a middle income group (a middle stable zero of the function) seems to have emerged in the mid-1970s (in that case, it was mainly comprised of Latin American countries), only to vanish in the early 1980s. A middle group emerged again in the late 1980s (the Asian Tigers), and it may be too early to tell whether or not the dynamics have changed yet again to only allow for the rich and poor clubs, with no middle-income group.[10] Those graphs heuristically support the two stylized facts: (i) a stable middle income group seems to emerge every now and then, and (ii) distribution dynamics change every so often, to destabilize that middle group. We now turn to formalizing those conclusions through a nonparametric estimation and hypothesis testing framework.

3. Stochastic Stability and the Convergence Club Hypothesis

We consider a Markov process on relative incomes, defined by an estimated transition function $P_n(x,A)$ from any point x to any set A, with density $t_n(x,y)$.[11] We shall define our stochastic stability conditions in terms of the properties of this transition

---

[10] To illustrate cases wherein our non-parametric estimate of $f(.)$ may give rise to a single stable zero, we show in Figure 4 the estimates for OECD countries’ income distribution, as well as those for cross-state income distributions in the U.S. However, even in those cases, we can see the emergence of more stable zeros in both cases. However, applying our tests to those cases is beyond the scope of this paper.

[11] Formally, for any $x \in \mathbb{R}$, $P_n(x,\cdot)$ is a probability measure on the Borel sigma-algebra $\mathfrak{B}(\mathbb{R})$, and for a given Borel set $A \in \mathfrak{B}(\mathbb{R})$, $P_n(\cdot,A)$ is a Borel measurable function.

density. We shall estimate the transition density over a compact set X via the kernel density estimator:

$$t_n(x,x') = q_n(x,x')/p_n(x), \qquad \text{where}$$
$$p_n(x) = \frac{1}{nh}\sum_{i=1}^{n} K(\frac{x-x_i}{h}) \quad \text{and} \quad q_n(x,x') = \frac{1}{nh}\sum_{i=1}^{n-1} K(\frac{x-x_i}{h^{1/2}})K(\frac{x'-x_{i+1}}{h^{1/2}})$$

are the corresponding non-parametric estimates of the marginal and joint densities, respectively. We then obtain our estimate of the conditional CDF of this transition as follows:

$$G_n(z \mid x) = \int_{-\infty}^{z} t_n(dx' \mid x)$$

Let the α-quantile of this CDF be denoted $\xi_n(\alpha, x)$. We assume that the median $\xi_n(0.5, x)$, defined as the smallest root of the equation $G_n(z \mid x) = 0.5$, is unique. We then investigate stochastic stability in terms of the function $f_n(x) = \xi_n(0.5, x) - x.$

Definition (Stochastic Stability)

A point x is said to be a stochastically stable zero of the function $f_n(.)$ if:

i) $f_n(x) = \xi_n(0.5, x) - x = 0,$

ii) $f_n(x') > f_n(x)$ for $x' \in (x-\varepsilon, x)$ and some $\varepsilon > 0,$ and

iii) $f_n(x') < f_n(x)$ for $x' \in (x, x+\delta)$ and some $\delta > 0.$

As noted in Section 2, our analogue of the two-clubs or “twin-peaks” hypothesis of Quah (1997) is thus represented as a function $f(.)$ with two stable zeros. The classical convergence hypothesis (e.g. as found for within-U.S. income distribution across states, or within the OECD block c.f. Johnson (2000)) is represented with a single stable zero. The existence of three stable zeros would question both the classical convergence hypothesis, as well as Quah’s “vanishing middle-income group” hypothesis. In the remainder of this section, we shall derive the statistical properties of the estimated $f_n(.)$, and thus obtain a formal test for the number of stable zeros.

Under the appropriate assumptions summarized in Appendix A, Roussas (1991) proved strong consistency of the estimated CDF $G_n(z \mid x)$ (Theorem 2.1 in Roussas (1991), p.446) as well as the estimated quantile $\xi_n(\alpha, x)$ (Theorem 2.2, ibid.). In addition, we can show under suitable assumptions that the estimated α-quantile $\xi_n(\alpha, x)$ has an asymptotic normal distribution as stated in the proposition 1.[12]

[12] The expression we derive for the asymptotic variance of the estimated quantile is derived using the techniques in Roussas (1991), and it is a correction for the variance term stipulated in Roussas (1969b). Assumptions A.1 through A.5 in Appendix A are relatively standard. The crucial assumption is that of

Proposition 1. Under suitable assumptions (see Appendix A for assumptions and proof)

$$(nh)^{1/2}[\xi_n(\alpha,x)-\xi(\alpha,x)] \xrightarrow{d} N(0,\tau^2(\xi,x)), \text{ where}$$

$$\tau^2(\xi,x)=t^{-2}(\xi\,|\,x)G\left(\xi\,|\,x\right)(1-G\left(\xi\,|\,x\right))\,p^{-1}(x)\int K^2(z)dz$$

Replacing the unknown quantities $\{t(\xi\,|\,x), G\left(\xi\,|\,x\right), p(x)\}$ with their consistent estimates, we can easily perform individual tests for the null hypothesis $H_0:\xi(0.5,y)=0$. In addition, once we show that the estimates at different points are asymptotically uncorrelated, we can easily construct uniform confidence bands around the function $f(.)$ based on the point-wise confidence bands. We can also perform simultaneous tests of multiple stable zeros, as shown in Section 4. For now, we conclude this section with two crucial results: uncorrelatedness of conditional median estimates for the same time period across relative incomes, and across time periods for the same relative income. Those two results will be crucial for performing tests regarding the number of zeros of the estimated stochastic $f(.)$. Proposition 2 establishes uncorrelatedness of the conditional median estimate at different points for the same period, thus allowing us to construct a simple Chi-squared test of multiple-zeros of the function by utilizing only estimated variances at the stipulated zeros. Since our cross-sectional sample is too small to yield sufficiently small estimated variances at the various points, we shall average estimated $f(.)$ across five time periods. The calculation of the estimated variance of the average is made simple due to Proposition 3, which shows that the estimated conditional median at the same point over different time periods is uncorrelated.

Proposition 2. Under suitable assumptions (see Appendix A for assumptions and proof)

$$\text{cov}[(\xi_n(\alpha,x)-\xi(\alpha,x)),(\xi_n(\alpha,y)-\xi(\alpha,y))] \xrightarrow{p} 0 \quad \text{for } x\neq y,\;\; x,y\in\mathbb{R}$$

Proposition 3. Under suitable assumptions (see Appendix A for assumptions and proof)

$$\text{cov}[(\xi_n^{\,t+1}(\alpha,x)-\xi^{t+1}(\alpha,x)),(\xi_n^{\,t}(\alpha,x)-\xi^{t}(\alpha,x))] \xrightarrow{p} 0 \quad \text{for } x\in\mathbb{R}$$

---

stationary-ergodicity/mixing of the true underlying transition kernel. While we question the stationarity of the transition kernel over the entire sample, asserting that different sub-periods were characterized by different transition kernels, we may still maintain that the transition kernel for each sub-period (were it to continue indefinitely) would indeed converge to a unique invariant (ergodic) distribution. When we estimate a transition kernel over a long period, we may think of the object of estimation as a mixture of the various ergodic transition kernels, which would itself be ergodic under appropriate technical assumptions (e.g. Markovian regime switching according to an ergodic probability transition matrix).

4. Hypothesis Testing Support for Stylized Facts

Before using our measure of stochastic stability to test the hypothesis of existence of a stable middle-income group of countries in various sub-periods of the sample, we present two test results to support the hypothesis of non-stationarity of income dynamics. Results of the first test are shown in Figure 5. This figure reports the results of Kolmogorov-Smirnov tests of equality of the conditional transition CDFs for each initial relative income level for the first and last years of our sample: 1961-2 and 1995-6, respectively.[13] As we can see, the *p*-values for this test are extremely small for almost all initial relative incomes, with the exception of a small and relatively unpopulated band of high-incomes. Consequently, we can conclude that for almost all initial relative incomes, the transition dynamics of countries with such relative incomes were significantly different in the different parts of our sample.

For comparison with KOS (2001) and other results in the literature, which focused on limiting distributions, we perform a second Kolmogorov-Smirnov test of equality of the estimated limiting income distributions of the estimated five year transitions for the first and last five-year periods in our sample: 1961-66 and 1991-96, respectively. The pdfs and CDFs of those limiting distributions are shown in Figure 6. The difference between those two limiting distributions is obvious, and it results in an extremely low p-value of 1.3799889e-009 for the Kolmogorov-Smirnov test of the equality of the two CDFs. Thus, Figures 5 and 6 further suggest that income distribution dynamics in our sample must be studied separately for the different sub-periods. In particular, we shall focus on the existence of a stochastically stable middle income group in various sub-periods. The existence of such a stochastically stable middle income group, together with a basin of attraction that is close to that of the poor countries, increases the probability of escape from a low-income poverty trap, as we shall see in Section 5. First, we turn in the remainder of this section to the formal tests of existence of a stochastically stable middle group in various sub-periods.

Utilizing the consistency and asymptotic normality results of Section 3, we can test formally for the existence of multiple stochastically stable income groups over various time-periods/epochs. Formally, our test will be constructed as follows:

---

[13] The classical Kolmogorov-Smirnov test with two empirical CDFs utilizes the central limit theorem for two i.i.d. samples, thus generating the K-S statistic as the supremum of a Brownian Bridge, and deriving the distribution of said maximum. In our case, the central limit theorems for the two estimated conditional transition CDFs $G_n(.|x)$ from the two cross-sectional samples (one in the early 60s and one in the early 90s) are based on the asymptotic normality result of Roussas (1991, Theorem 2.3, p.447). Covariances between the two estimated CDFs for the early 1960s and 1990s are assumed negligible under the maintained hypothesis of an overall mixing DGP.

$$H_0 : (\xi_n(0.5,x)-x)|_{c_1}=0 \text{ and } (\xi_n(0.5,x)-x)|_{c_2}=0 \text{ and} \cdots \text{and } (\xi_n(0.5,x)-x)|_{c_l}=0$$

vs.

$$H_1 : (\xi_n(0.5,x)-x)|_{c_1}\neq 0 \text{ or } (\xi_n(0.5,x)-x)|_{c_2}\neq 0 \text{ or} \cdots \text{or } (\xi_n(0.5,x)-x)|_{c_l}\neq 0,$$

where $\xi_n(0.5,x)$ is an estimated conditional median of income at t+1, and c_s, s=1,2,…,l, is the list of points at which our function is equal to zero under the null hypothesis.[14] For instance, to have three stable zeros of our function $f_n(x)=\xi_n(0.5,x)-x$, we would need five zeros of the function, as shown in Figure 2. If we have only two stable zeros (our dynamic version of the twin-peaks hypothesis), then we should reject the null hypothesis of five zeros, in favor of only three zeros (two of which would be stable: representing the rich and poor clubs) or less. Given our interest in testing the three clubs hypothesis (including a middle-income group), we conducted the test at 5 candidate zeros of our function. If we were interested only in the two-clubs vs. global convergence hypotheses, we would have tested the null of three zeros vs. the alternative of only one. The stylized facts explored in Section 2 prompted us to conduct the formal test for 5 zeros, three of which would be stable.

Since the quantile estimator $\xi_n(\alpha,x)$ is asymptotically normally distributed, with the covariance matrix derived in Section 3, we can construct our test statistic as a quadratic form with an asymptotic χ^2 distribution under the null hypothesis. Formally, our test statistic is:

$$(\xi_n(0.5,c_1)-c_1,\ \xi_n(0.5,c_2)-c_2,\cdots,\xi_n(0.5,c_l)-c_l)\begin{pmatrix}\tau_1^2 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \tau_l^2\end{pmatrix}^{-1}\begin{pmatrix}\xi_n(0.5,c_1)-c_1\\ \xi_n(0.5,c_2)-c_2\\ \vdots\\ \xi_n(0.5,c_l)-c_l\end{pmatrix}$$

$$=\sum_{s=1}^{l}\left\{\frac{\xi_n(0.5,c_s)-c_s}{\tau_s}\right\}^2 \xrightarrow[\text{(under } H_0)]{d} \chi^2(l).$$

In application, the standard errors τ_s for annual transition estimates were too large. Consequently, we utilized the result in Proposition 3, to average estimates over five year periods, thus obtaining reasonably small confidence intervals on $\xi_n(0.5,c)-c$.

Tables 4-10 show the test results for seven five-year periods, starting with 1961-66, and ending with 1991-96. For instance, Table 4 shows that the test statistic for the period 1961-66 exceeded its 5% critical value. Hence, we reject the null hypothesis of three stable zeros, in favor of the fewer-clubs (two: rich and poor) hypothesis. The same result is also obtained for the period 1966-71. However, for the next two five year

[14] The five points at which the tests were conducted were chosen for each five-year period were selected based on the zeros of the function based on one-year transitions (see Figure 6).

periods: 1971-76 and 1976-81, we fail to reject the null hypothesis of three stochastically stable zeros. In other words, the test results in Tables 6 and 7 confirm the ocular reading of the graphed function to the right of Table 7, that we have a stochastically stable middle-income group in the mid-to-late 1970s. The test in Table 8 rejects the null hypothesis of three groups at the 5% significance level for the period 1981-86. Then, a stable middle-income group re-emerges in the late 1980s and early 1990s, resulting in failure to reject the null hypothesis for those two periods. In other words, stochastically stable middle income groups seem to have appeared in the mid 1970s (mostly Latin American countries), and then to have disappeared in the early 1980s as most of those countries drifted downward in relative per capita incomes. Then, another stable middle-income group seemed to have arisen in the late 1980s and early 1990s (mostly Asian Tigers). It may be too early to determine whether or not this group has been destabilized by the financial difficulties most emerging markets suffered during the late 1990s.

Those formal test results support the stylized facts of Section 2, which were based on graphical inspection of our measure of stability of various income groups. They also suggest that framing the problem in terms of "convergence"-type hypotheses of any kind may obscure some of the most interesting dynamics. In particular, it would be very useful to model the changes in stochastic transitions themselves, which give rise to the periodic appearance and disappearance of a stable middle-income group. For example, casual ocular inspection of Figure 7, which shows all the zeros of the function $f_n(.)$ estimated for each 1-year transitions, suggests that a middle-income group tends to emerge as the gap between the rich and poor clubs gets wider, and to disappear as this gap shrinks in later periods. However, modeling this higher-level dynamic would require a different methodology, and thus remains beyond the scope of this paper.

5. Economic Implications and Concluding Remarks

Quah (2001) noted that the major weakness in the distributional dynamics approach to income distributions literature pertains to the areas of formal statistical testing and inference. Bianchi (1997) and KOS (2001) provided valuable early attempts to fill this gap in the literature. However, the first approach remained primarily static (test of multi-modality of estimated cross-country income distribution density at a given point in time), while the second focused only on the asymptotic (ergodic) distribution, to the exclusion of short-term dynamics. As also noted by Quah (2001), this emphasis on the limiting distributions of estimated probability transition matrices is not particularly insightful, since classical convergence may only be attainable in this framework after centuries or millennia of increased inequality between the rich and poor.

The two stylized facts we established heuristically in Section 2, and supported with hypothesis tests in Section 4, suggest strongly that distributional dynamics are non-stationary over our sample of 35 years. Examining transition dynamics over 5-year periods, we detect strong evidence for periodic appearance and disappearance of a stochastically stable middle-income group. The significance of this middle income group is the fact that it has a "stochastic basin of attraction" that should make it easier for poor countries to escape the poverty trap. For instance, consider the schematic diagram in Figure 2. In this diagram with three stable zeros, the "stochastic basin of attraction" for the poor club extends only up to C2, the basin of attraction for the middle club extends from C2 to C4, and the basin of attraction for the rich group extends from C4 upwards. In periods with no middle group (the curve stays below the *x*-axis, and only crosses at C4), there would be a large basin of attraction for the poor club extending up to C4, and a basin of attraction for the rich club above that point. For a poor country to escape the poverty trap (interpreted here as the stochastic basin of attraction for the poor club), it has to sustain a "big push" that transfers it to another basin of attraction. The existence of a stable middle group brings such an alternative basin of attraction closer, and thus makes it more likely for any given country to escape the poverty trap. This is illustrated in Figure 8, where we plot the probability of a country escaping from the poor club (a relative income close to the first stable zero) to the middle income range (a relative income in the middle of the range of our observations), for different years. Using the estimated one-year transition kernels, we can see that the probability of escape asymptotes to a low invariant probability very quickly for the years 1961-62, 1971-72, and 1981-82, during which we are finding no stable middle group. In contrast, we find that the probability of escape from the poverty trap continues to grow for the years 1975-76 and 1991-92, during which such a stable middle group was present.

The sample of 35 years is too short to fully understand the nature of higher-order dynamics. In Figure 7, where we traced the locus of zeros of the conditional median relative income less initial relative income, we can see that the appearance of a stable middle group coincides with a widening of the gap between the rich and the poor, and its disappearance coincides with a narrowing of this gap. Indeed, Paap and Dijk (1998) have also documented this stylized fact. However, understanding the causal underpinnings of this constant conjunction, and the economic conditions that are conducive for the appearance of a stable middle-income group, requires deeper economic modeling of the higher level dynamics. Understanding those higher level dynamics appears to be a more fruitful than defending globalization on the basis of "convergence" or condemning it on the basis of "divergence", when both hypotheses in fact appear to be ill-posed.

Appendix A

We use the nonparametric estimates of the marginal and joint pdfs following Roussas (1969a, 1969b, 1991):

$$p_n(x) = (nh)^{-1} \sum_{j=1}^{n} K((x - X_j)h^{-1})$$
$$q_n(x, x') = (nh)^{-1} \sum_{j=1}^{n-1} K((x - X_j)h^{-1/2}) K((x' - X_{j+1})h^{-1/2}).$$

The natural nonparametric estimate of the transition pdf is therefore:

$$t_n(x' \mid x) = q_n(x, x') / p_n(x).$$

This in turn gives rise to a reasonable estimate of the transition CDF:

$$G_n(z \mid x) = \int_{-\infty}^{z} t_n(dx' \mid x).$$

Roussas (1969a, 1969b, 1991) makes the following regularity assumptions for proving consistency and asymptotic normality of our statistics of interest:

A.1.

(i) The real-valued random variables X_n, $n \geq 1$, are defined on a probability space $(\Omega, \mathcal{A}, P)$ and constitute a strictly stationary Markov process with a single ergodic set and no cyclically moving subsets.

(ii) The marginal and joint distributions of X_1 and (X_1, X_2) have continuous pdfs with respect to appropriate Lebesgue measures: $p(.)$ and $q(.,.)$, respectively.

(iii) The process $\{X_n\}$ is ρ-mixing with maximal correlation $\rho(n) = O(n^{-\upsilon}), \upsilon > 1$.

A.2. K is a pdf defined on $\mathbb{R}$ such that:

(i) K is bounded.

(ii) $|x| K(x) \to 0$ as $|x| \to \infty$.

(iii) $\int xK(x)dx = 0$ and $\int x^2 K(x)dx < \infty$.

A.3. $h = h(n)$ stands for a sequence of real numbers such that as $n \uparrow \infty$:

(i) $0 < h \to 0$.

(ii) For some arbitrarily large $l \geq 2, \sum_{n=1}^{\infty} (n^{1/2}h)^{-l} < \infty$.

(iii) $nh^5 \to 0$.

A.4. For $\alpha \in (0,1)$ and $x \in \mathbb{R}$, the one-step transition CDF $G(. \mid x)$, has a unique α-quantile $\xi(\alpha, x)$.

A.5. (i) $p(.)$ has a continuous and bounded second order derivative.

(ii) The pdf $q(.,.)$ has continuous second order partial derivatives, denoted $q''_{ij}(.,.)$, i,j=1,2, such that $\int \left| q''_{ij}(x,y) \right| dy \leq C$ (independent of $x \in \mathbb{R}$).

(iii) $p(x) \leq C, x \in \mathbb{R}$, and $\left| q_{1,i}(x,y) - p(x)p(y) \right| \leq C, i \geq 2, \ x, y \in \mathbb{R}$.

We now list the major consistency and asymptotic normality results for the estimators of the transition CDF and α-quantile, as provided by Roussas (1969a,b, 1991):

THEOREM 1. (Roussas 1991, Theorem 2.1)
Under A.1; A.2 (i), (ii); and A.3 (i), (ii):

$$\sup\{|G_n(z \mid x) - G(z \mid x)|; z \in \mathbb{R}\} \to 0 \text{ a.s., } x \in \mathbb{R}$$

THEOREM 2. (Roussas 1991, Theorem 2.2)
Under A.1; A.2 (i), (ii); A.3 (i), (ii): and A.4, let $\xi_n(\alpha, x)$ denote the α-quantile of $G_n(z \mid x)$, $0 < \alpha < 1$, then:

$$\xi_n(\alpha, x) \to \xi(\alpha, x), \text{ a.s., } x \in \mathbb{R}$$

We now list the three propositions in Section 3 of this paper. Proposition 1 is a slight modification of the asymptotic normality result in Roussas (1969b). The expression for the variance term τ^2 in that paper is not correct. For completeness, we use the results and methods of Roussas (1991) to derive the correct expression for that variance term. In addition, those terminology and methods of Roussas (1991) will be used in proving Propositions 2 and 3, which assure us that covariance terms vanish asymptotically, thus allowing us to compute our quadratic form test statistic with great ease.

PROPOSITION 1. Under A.1-A.5:

$$(nh)^{1/2}[\xi_n(\alpha, x) - \xi(\alpha, x)] \xrightarrow{d} N(0, \tau^2(\xi, x)),$$
$$\text{where } \tau^2(\xi, x) = t^{-2}(\xi \mid x) G(\xi \mid x)(1 - G(\xi \mid x)) p^{-1}(x) \int K^2(z) dz$$

PROPOSITION 2. Under A.1-A.5:

$$\operatorname{cov}(\xi_n(\alpha, x) - \xi(\alpha, x), \xi_n(\alpha, y) - \xi(\alpha, y)) \xrightarrow{p} 0$$
$$\text{for } x \neq y, \quad x, y \in \mathbb{R}$$

PROPOSITION 3. Under A.1-A.5:

$$\operatorname{cov}(\xi_n^{t+1}(\alpha, x) - \xi^{t+1}(\alpha, x), \xi_n^{t}(\alpha, x) - \xi^{t}(\alpha, x)) \xrightarrow{p} 0$$
$$\text{for } x \in \mathbb{R}$$

PROOF of PROPOSITION 1

We first sketch the proof methodology of Roussas (1969a, 1969b, 1991) to derive the general variance term. The proof utilizes a Taylor expansion:

$$G_n(\xi_n \mid x) = G_n(\xi \mid x) + (\xi_n - \xi) t_n(\xi_n^* \mid x),$$

Where ξ_n is the α-quantile of G_n, and ξ is the α-quantile of G. Consequently, by construction, $G_n(\xi_n \mid x) = \alpha = G(\xi \mid x)$. Combining the two expressions, we get:

$$(nh)^{1/2}(\xi_n - \xi) = -(nh)^{1/2}[G_n(\xi \mid x) - G(\xi \mid x)]t_n^{-1}(\xi_n^* \mid x)$$

where ξ_n^* is a random variable whose values lie between ξ_n and ξ. Roussas (1969b, Lemma 5.1) showed that the last term converges in probability: $t_n(\xi_n^* \mid x) \xrightarrow{p} t(\xi \mid x), x \in \mathbb{R}$, and then proceeded to prove that $(nh)^{1/2}[G_n(\xi \mid x) - G(\xi \mid x)]$ has an asymptotically normal distribution. This is accomplished by writing:

$$G_n(\xi \mid x) = \sum_{j=1}^{n-1} L_n^*(Y_j) / \sum_{j=1}^{n} L_n(X_j), \text{ where}$$
$$L_n(X_j) = K((x - X_j)h^{-1}),$$
$$L_n^*(Y_j) = K((x - X_j)h^{-1/2}) \int_{-\infty}^{\xi} K((x' - X_{j+1})h^{-1/2})dx' \equiv K((x - X_j)h^{-1/2})\bar{K}((\xi(x) - X_{j+1})h^{-1/2}),$$
$$Y_j = (X_j, X_{j+1}),\ j=1,2\ldots,.n\text{-}1$$

Adding and subtracting $v_n = -[EL_n^*(Y_1)][EL_n^*(X_1)]^{-1}$

$$(nh)^{1/2}[G_n(\xi \mid x) - G(\xi \mid x)] = (nh)^{1/2}[(G_n(\xi \mid x) + v_n) - (G(\xi \mid x) + v_n)]$$

The second term converges in probability to zero (Roussas 1969b, Lemma 5.3), and the first term can be rewritten as:

$$(nh)^{1/2}[G_n(\xi_n \mid x) + v_n] = [(nh)^{-1} \sum_{j=1}^{n} L_n(X_j)]^{-1} (nh)^{-1/2} \sum_{j=1}^{n-1} [\varphi_n(Y_j) - E\varphi_n(Y_j)],$$
$$\text{where} \quad \varphi_n(Y_j) = L_n^*(Y_j) + v_n L_n(X_j)$$

Roussas (1991, Lemma 3.1) further proves that $(nh)^{-1} \sum_{j=1}^{n} L_n(X_j) \xrightarrow{p} p(x),$ a.s. $x \in \mathbb{R}$ and proceeds to prove asymptotic normality for the remaining term:

$$(nh)^{-1/2} \sum_{j=1}^{n-1} [\varphi_n(Y_j) - E\varphi_n(Y_j)] \xrightarrow{d} N(0, \sigma_0(\xi, x))$$
$$\text{where} \qquad \sigma_0(\xi, x) = \frac{q(x,\xi)[p(x) - q(x,\xi)] \int K^2(z)dz}{p(x)}$$

Returning to the α-quantile expansion:

$$(nh)^{1/2}(\xi_n - \xi) = -\{(nh)^{1/2}[G_n(\xi \mid x) + v_n] - (nh)^{1/2}[v_n + G(\xi \mid x)]\} t_n^{-1}(\xi_n^* \mid x),$$

it is clear that the asymptotic variance term is:

$$\tau^2(\xi, x) = t^{-2}(\xi \mid x)\sigma_0(\xi, x) p^{-2}(x) = t^{-2}(\xi \mid x) \frac{q(x,\xi)[p(x) - q(x,\xi)] \int K^2(z)dz}{p(x)} p^{-2}(x)$$
$$= t^{-2}(\xi \mid x) G(\xi \mid x)[1 - G(\xi \mid x)] \int K^2(z)dz\, p^{-1}(x)$$

PROOF of PROPOSITION 2:

We begin with the familiar Taylor expansion (thereafter, we suppress dependence on α):

$$(nh)^{1/2}(\xi_n(\alpha,x)-\xi(\alpha,x))\cdot(nh)^{1/2}(\xi_n(\alpha,y)-\xi(\alpha,y))$$
$$=\{(nh)^{1/2}[G_n(\xi(x)\,|\,x)-G(\xi(x)\,|\,x)]\}\cdot\{(nh)^{1/2}[G_n(\xi(y)\,|\,y)-G(\xi(y)\,|\,y)]\}t_n^{-1}(\xi_n^*\,|\,x)\cdot t_n^{-1}(\xi_n^*\,|\,y)$$

The last two terms converge in probability to $t^{-1}(\xi\,|\,x),$ and $t^{-1}(\xi\,|\,y)$, respectively. Therefore, it only remains to show that the expectation of the term inside the brackets converges in probability to zero.

Using Roussas's result that $(nh)^{1/2}[G(\xi\,|\,x)+v_n]\xrightarrow{p}0,$ and augmenting earlier notation to denote the previously used (L,L^*,ϕ) by (L^x,L^{*x},ϕ^x) , and the corresponding terms with Kernels centered around y by (L^y,L^{*y},ϕ^y) , we get:

$$\operatorname*{plim}_{n\uparrow\infty}\{(nh)[G_n(\xi(x)\,|\,x)-G(\xi(x)\,|\,x)]\cdot[G_n(\xi(y)\,|\,y)-G(\xi(y)\,|\,y)]\}$$
$$=\operatorname*{plim}_{n\uparrow\infty}\{(nh)[G_n(\xi(x)\,|\,x)+v_n(x)]\cdot[G_n(\xi(y)\,|\,y)+v_n(y)]\}$$
$$=\operatorname*{plim}_{n\uparrow\infty}\{[(nh)^{-1}\sum_{j=1}^{n}L_n^x(X_j)]^{-1}(nh)^{-1/2}\sum_{j=1}^{n-1}[\varphi_n^x(Y_j)-E\varphi_n^x(Y_j)]\times$$
$$[(nh)^{-1}\sum_{j=1}^{n}L_n^y(X_j)]^{-1}(nh)^{-1/2}\sum_{j=1}^{n-1}[\varphi_n^y(Y_j)-E\varphi_n^y(Y_j)]\}$$
$$=p^{-1}(x)p^{-1}(y)\{(nh)^{-1/2}\sum_{j=1}^{n-1}[\varphi_n^x(Y_j)-E\varphi_n^x(Y_j)]\cdot(nh)^{-1/2}\sum_{j=1}^{n-1}[\varphi_n^y(Y_j)-E\varphi_n^y(Y_j)]\}.$$

We now need to show that the expectation of the term in brackets converges in probability to zero. Let

$$\Phi_n(x,y)\equiv(nh)^{-1/2}\sum_{j=1}^{n-1}[\varphi_n^x(Y_j)-E\varphi_n^x(Y_j)]\cdot(nh)^{-1/2}\sum_{j=1}^{n-1}[\varphi_n^y(Y_j)-E\varphi_n^y(Y_j)]$$
$$=(nh)^{-1}\sum_{j=1}^{n-1}[L_n^{*x}(Y_j)+v_n(x)L_n^x(X_j)-E(L_n^{*x}(Y_j)+v_n(x)L_n^x(X_j))]\times$$
$$\sum_{j=1}^{n-1}[L_n^{*y}(Y_j)+v_n(y)L_n^y(X_j)-E(L_n^{*y}(Y_j)+v_n(y)L_n^y(X_j))]$$
$$=(nh)^{-1}\sum_{j=1}^{n-1}[(L_n^{*x}(Y_j)-EL_n^{*x}(Y_j))+v_n(x)(L_n^x(X_j)-EL_n^x(X_j))]\times$$
$$\sum_{j=1}^{n-1}[(L_n^{*y}(Y_j)-EL_n^{*y}(Y_j))+v_n(y)(L_n^y(X_j)-EL_n^y(X_j))]\ .$$

Now taking the expectation of $\Phi_n(x,y)$:

$$
\begin{aligned}
E[\Phi_n(x,y)] &= \frac{v_n(x)\cdot v_n(y)}{h}\operatorname{cov}\left[K\left(\frac{x-X_1}{h}\right),K\left(\frac{y-X_1}{h}\right)\right]+\\
&\frac{1}{h}\operatorname{cov}\left[K\left(\frac{x-X_1}{h}\right)\overline{K}\left(\frac{\xi(x)-X_2}{h}\right),K\left(\frac{y-X_1}{h}\right)\overline{K}\left(\frac{\xi(y)-X_2}{h}\right)\right]+\\
&\frac{v_n(x)}{h}\operatorname{cov}\left[K\left(\frac{x-X_1}{h}\right),K\left(\frac{y-X_1}{h}\right)\overline{K}\left(\frac{\xi(y)-X_2}{h}\right)\right]+\\
&\frac{v_n(y)}{h}\operatorname{cov}\left[K\left(\frac{y-X_1}{h}\right),K\left(\frac{x-X_1}{h}\right)\overline{K}\left(\frac{\xi(x)-X_2}{h}\right)\right]+\\
&\frac{1}{nh}\sum_{1\le i\le j\le n-1}\operatorname{cov}\left[K\left(\frac{x-X_i}{h}\right)\overline{K}\left(\frac{\xi(x)-X_{i+1}}{h}\right),K\left(\frac{y-X_j}{h}\right)\overline{K}\left(\frac{\xi(y)-X_{j+1}}{h}\right)\right]+\\
&\frac{1}{nh}\sum_{1\le i\le j\le n-1}\operatorname{cov}\left[K\left(\frac{x-X_j}{h}\right)\overline{K}\left(\frac{\xi(x)-X_{j+1}}{h}\right),K\left(\frac{y-X_i}{h}\right)\overline{K}\left(\frac{\xi(y)-X_{i+1}}{h}\right)\right]+\\
&\frac{v_n(x)}{nh}\sum_{1\le i\le j\le n-1}\operatorname{cov}\left[K\left(\frac{x-X_i}{h}\right),K\left(\frac{y-X_j}{h}\right)\overline{K}\left(\frac{\xi(y)-X_{j+1}}{h}\right)\right]+\\
&\frac{v_n(x)}{nh}\sum_{1\le i\le j\le n-1}\operatorname{cov}\left[K\left(\frac{x-X_j}{h}\right),K\left(\frac{y-X_i}{h}\right)\overline{K}\left(\frac{\xi(y)-X_{i+1}}{h}\right)\right]+\\
&\frac{v_n(y)}{nh}\sum_{1\le i\le j\le n-1}\operatorname{cov}\left[K\left(\frac{x-X_i}{h}\right)\overline{K}\left(\frac{\xi(x)-X_{i+1}}{h}\right),K\left(\frac{y-X_j}{h}\right)\right]+\\
&\frac{v_n(y)}{nh}\sum_{1\le i\le j\le n-1}\operatorname{cov}\left[K\left(\frac{x-X_j}{h}\right)\overline{K}\left(\frac{\xi(x)-X_{j+1}}{h}\right),K\left(\frac{y-X_i}{h}\right)\right]+\\
&\frac{v_n(x)\cdot v_n(y)}{nh}\sum_{1\le i\le j\le n-1}\operatorname{cov}\left[K\left(\frac{x-X_i}{h}\right),K\left(\frac{y-X_j}{h}\right)\right]+\\
&\frac{v_n(x)\cdot v_n(y)}{nh}\sum_{1\le i\le j\le n-1}\operatorname{cov}\left[K\left(\frac{x-X_j}{h}\right),K\left(\frac{y-X_i}{h}\right)\right]
\end{aligned}
$$

In fact, under our assumptions on the kernel K (the tails of which converge to zero), and for any $x \neq y$, as $n \uparrow \infty$, $h \downarrow 0$, and the first covariance term will vanish:

$$
\begin{aligned}
&\frac{1}{h}\operatorname{cov}\left[K\left(\frac{x-X_1}{h}\right),K\left(\frac{y-X_1}{h}\right)\right]\\
&=\frac{1}{h}\left\{E\left[K\left(\frac{x-X_1}{h}\right)K\left(\frac{y-X_1}{h}\right)\right]-E\left[K\left(\frac{x-X_1}{h}\right)\right]\cdot E\left[K\left(\frac{y-X_1}{h}\right)\right]\right\}\\
&=\frac{1}{h}\int K(u)K\left(\frac{y-X_1}{h}\right)p(x-hu)(hdu)-\frac{1}{h}\left[\int K(u)\,p(x-hu)(hdu)\cdot\int K(v)\,p(y-hv)(hdv)\right]\\
&=\int K(u)K\left(\frac{y-x+hu}{h}\right)p(x-hu)du-h\left[\int K(u)\,p(x-hu)du\int K(v)\,p(y-hv)dv\right]\\
&\longrightarrow 0.
\end{aligned}
$$

Similarly, the third and fourth terms can be shown to vanish asymptotically. The third covariance term contains two parts. The first component:

$$
\begin{aligned}
&\frac{1}{h}\left\{E\left[K\left(\frac{y-X_1}{h}\right)\overline{K}\left(\frac{\xi(y)-X_2}{h}\right)\right]\cdot E\left[K\left(\frac{x-X_1}{h}\right)\right]\right\}\\
&\leq\frac{1}{h}\left\{E\left[K\left(\frac{y-X_1}{h}\right)\right]\cdot E\left[K\left(\frac{x-X_1}{h}\right)\right]\right\}\\
&=h\left[\int K(u)\,p(y-hu)du\cdot\int K(v)\,p(x-hv)dv\right]\to 0.
\end{aligned}
$$

The second component of the second covariance term also vanishes:

$$
\begin{aligned}
&\frac{1}{h}E\left[K\left(\frac{y-X_1}{h}\right)\overline{K}\left(\frac{\xi(y)-X_2}{h}\right)K\left(\frac{x-X_1}{h}\right)\right]\\
&=\int K(u)K\left(\frac{x-X_1}{h}\right)\left[\int\overline{K}\left(\frac{\xi(y)-t}{h}\right)q(y-hu,t)dt\right]du\\
&\to 0
\end{aligned}
$$

The fourth covariance term can similarly be decomposed into two parts, each of which vanishes in the same manner.

Finally, all remaining covariance terms vanish asymptotically by Lemmas 5.4 through 5.6 in Roussas (1991).

PROOF of PROPOSITION 3:

The proof of this proposition is very similar to that of Proposition 2. We begin with the Taylor expansion:

$$
\begin{aligned}
&(nh)^{1/2}(\xi_n^{\,t+1}(x,\alpha)-\xi^{t+1}(x,\alpha))\cdot(nh)^{1/2}(\xi_n^{\,t}(x,\alpha)-\xi^{t}(x,\alpha))\\
&=\{(nh)^{1/2}[G_n(\xi^{t+1}(x)\,|\,x)-G(\xi^{t+1}(x)\,|\,x)]\}\cdot\{(nh)^{1/2}[G_n(\xi^{t}(x)\,|\,x)-G(\xi^{t}(x)\,|\,x)]\}\\
&\quad\cdot t_n^{-1}(\xi_n^{*t+1}\,|\,x)\cdot t_n^{\,-1}(\xi_n^{*t}\,|\,x_t)
\end{aligned}
$$

Following the same procedure utilized in Proposition 2, we get:

$$
\begin{aligned}
&\operatorname*{plim}_{n\uparrow\infty}\ \{(nh)[G_n(\xi^t(x)\,|\,x)-G(\xi^t(x)\,|\,x)]\cdot[G_n(\xi^{t+1}(x)\,|\,x)-G(\xi^{t+1}(x)\,|\,x)]\}\\
&=\operatorname*{plim}_{n\uparrow\infty}\ \{(nh)[G_n(\xi^t(x)\,|\,x)+v_n^t(x)]\cdot[G_n(\xi^{t+1}(x)\,|\,x)+v_n^{t+1}(x)]\}\\
&=\operatorname*{plim}_{n\uparrow\infty}\ \{[(nh)^{-1}\textstyle\sum_{j=1}^{n}L_n^t(X_j)]^{-1}(nh)^{-1/2}\sum_{j=1}^{n-1}[\varphi_n^t(Y_j)-E\varphi_n^t(Y_j)]\cdot\\
&\qquad\qquad[(nh)^{-1}\textstyle\sum_{j=1}^{n}L_n^{t+1}(X_j)]^{-1}(nh)^{-1/2}\sum_{j=1}^{n-1}[\varphi_n^{t+1}(Y_j)-E\varphi_n^{t+1}(Y_j)]\}\\
&=\frac{1}{p(x)\cdot p(x)}\{(nh)^{-1/2}\textstyle\sum_{j=1}^{n-1}[\varphi_n^t(Y_j)-E\varphi_n^t(Y_j)]\cdot(nh)^{-1/2}\sum_{j=1}^{n-1}[\varphi_n^{t+1}(Y_j)-E\varphi_n^{t+1}(Y_j)]\},
\end{aligned}
$$

where

$$
\begin{aligned}
&\varphi_n^t(Y_j)=L_n^{*t}(Y_j)+v_n^t(x)L_n^t(X_j)\ ,\quad \varphi_n^{t+1}(Y_j)=L_n^{*t+1}(Y_j)+v_n^{t+1}(x)L_n^{t+1}(X_j)\\
&L_n^t(X_j)=K((x-X_j)h^{-1}),\ L_n^{*t}(Y_j)=K((x-X_j)h^{-1/2})\bar{K}((\xi(x)-X_{j+1})h^{-1/2})\\
&L_n^{t+1}(X_j)=K((x-X_{j+1})h^{-1}),\ L_n^{*t+1}(Y_{j+1})=K((x-X_{j+1})h^{-1/2})\bar{K}((\xi(x)-X_{j+2})h^{-1/2})\\
&v_n^t(x)=-[EL_n^{*t}(Y_1)][EL_n^{*t}(X_1)]^{-1},v_n^{t+1}(x)=-[EL_n^{*t+1}(Y_1)][EL_n^{*t+1}(X_1)]^{-1}
\end{aligned}
$$

Thus, it suffices to get the expectation term expression above. Again, for convenience, define

$$
\begin{aligned}
&\Phi_n(t,t+1)\\
&\equiv(nh)^{-1/2}\textstyle\sum_{j=1}^{n-1}[\varphi_n^t(Y_j)-E\varphi_n^t(Y_j)]\times(nh)^{-1/2}\sum_{j=1}^{n-1}[\varphi_n^{t+1}(Y_j)-E\varphi_n^{t+1}(Y_j)]\\
&=(nh)^{-1}\textstyle\sum_{j=1}^{n-1}[L_n^{*t}(Y_j)+v_n^t(x)L_n^t(X_j)-E(L_n^{*t}(Y_j)+v_n^t(x)L_n^t(X_j))]\times\\
&\qquad\qquad\textstyle\sum_{j=1}^{n-1}[L_n^{*t+1}(Y_j)+v_n^{t+1}(x)L_n^{t+1}(X_j)-E(L_n^{*t+1}(Y_j)+v_n^{t+1}(x)L_n^{t+1}(X_j))]\\
&=(nh)^{-1}\textstyle\sum_{j=1}^{n-1}[(L_n^{*t}(Y_j)-EL_n^{*t}(Y_j))+v_n^t(x)(L_n^t(X_j)-EL_n^t(X_j))]\times\\
&\qquad\qquad\textstyle\sum_{j=1}^{n-1}[(L_n^{*t+1}(Y_j)-EL_n^{*t+1}(Y_j))+v_n^{t+1}(x)(L_n^{t+1}(X_j)-EL_n^{t+1}(X_j))]
\end{aligned}
$$

Taking expectations, we get

$$
\begin{aligned}
E[\Phi_n(t,t+1)] = {} & \frac{v_n^t(x)\cdot v_n^{t+1}(x)}{h}\operatorname{cov}\left[K\left(\frac{x-X_1}{h}\right),K\left(\frac{x-X_2}{h}\right)\right]+ \\
& \frac{1}{h}\operatorname{cov}\left[K\left(\frac{x-X_1}{h}\right)\overline{K}\left(\frac{\xi^t(x)-X_2}{h}\right),K\left(\frac{x-X_2}{h}\right)\overline{K}\left(\frac{\xi^{t+1}(x)-X_3}{h}\right)\right]+ \\
& \frac{v_n^t(x)}{h}\operatorname{cov}\left[K\left(\frac{x-X_1}{h}\right),K\left(\frac{x-X_2}{h}\right)\overline{K}\left(\frac{\xi^{t+1}(x)-X_3}{h}\right)\right]+ \\
& \frac{v_n^{t+1}(x)}{h}\operatorname{cov}\left[K\left(\frac{x-X_2}{h}\right),K\left(\frac{x-X_1}{h}\right)\overline{K}\left(\frac{\xi^t(x)-X_2}{h}\right)\right]+ \\
& \frac{1}{nh}\sum_{1\le i\le j\le n-1}\operatorname{cov}\left[K\left(\frac{x-X_i}{h}\right)\overline{K}\left(\frac{\xi^t(x)-X_{i+1}}{h}\right),K\left(\frac{x-X_{j+1}}{h}\right)\overline{K}\left(\frac{\xi^{t+1}(x)-X_{j+2}}{h}\right)\right]+ \\
& \frac{1}{nh}\sum_{1\le i\le j\le n-1}\operatorname{cov}\left[K\left(\frac{x-X_j}{h}\right)\overline{K}\left(\frac{\xi^t(x)-X_{j+1}}{h}\right),K\left(\frac{x-X_{i+1}}{h}\right)\overline{K}\left(\frac{\xi^{t+1}(x)-X_{i+2}}{h}\right)\right]+ \\
& \frac{v_n^t(x)}{nh}\sum_{1\le i\le j\le n-1}\operatorname{cov}\left[K\left(\frac{x-X_i}{h}\right),K\left(\frac{x-X_{j+1}}{h}\right)\overline{K}\left(\frac{\xi^{t+1}(x)-X_{j+2}}{h}\right)\right]+ \\
& \frac{v_n^t(x)}{nh}\sum_{1\le i\le j\le n-1}\operatorname{cov}\left[K\left(\frac{x-X_j}{h}\right),K\left(\frac{x-X_{i+1}}{h}\right)\overline{K}\left(\frac{\xi^{t+1}(x)-X_{i+2}}{h}\right)\right]+ \\
& \frac{v_n^{t+1}(x)}{nh}\sum_{1\le i\le j\le n-1}\operatorname{cov}\left[K\left(\frac{x-X_i}{h}\right)\overline{K}\left(\frac{\xi^t(x)-X_{i+1}}{h}\right),K\left(\frac{x-X_{j+1}}{h}\right)\right]+ \\
& \frac{v_n^{t+1}(x)}{nh}\sum_{1\le i\le j\le n-1}\operatorname{cov}\left[K\left(\frac{x-X_j}{h}\right)\overline{K}\left(\frac{\xi^t(x)-X_{j+1}}{h}\right),K\left(\frac{x-X_{i+1}}{h}\right)\right]+ \\
& \frac{v_n^t(x)\cdot v_n^{t+1}(x)}{nh}\sum_{1\le i\le j\le n-1}\operatorname{cov}\left[K\left(\frac{x-X_i}{h}\right),K\left(\frac{x-X_{j+1}}{h}\right)\right]+ \\
& \frac{v_n^t(x)\cdot v_n^{t+1}(x)}{nh}\sum_{1\le i\le j\le n-1}\operatorname{cov}\left[K\left(\frac{x-X_j}{h}\right),K\left(\frac{x-X_{i+1}}{h}\right)\right]
\end{aligned}
$$

Utilizing the same logic used in Proposition 2, and using Lemmas 5.4 through 5.6 of Roussas (1991), we can see that all covariance terms vanish asymptotically.

Appendix B

A limiting-distribution version of Bianchi (1997)'s test: more support for non-stationarity

The previous literature focused primarily on tests of bi-modality vs. uni-modality of a unique invariant distribution. Two formal tests were developed in that earlier literature: Bianchi (1997) utilized a formal test of multi-modality of kernel-estimated cross-country income distribution densities at different points in time. KOS (2001) performed a more dynamic likelihood ratio test of multi-modality of the ergodic (limiting) distribution of the estimated 5-year transition probability matrix.[15]

Bianchi (1997) employed Silverman's "critical bandwidth" as a test statistics for the hypothesis of multi-modality. The *critical bandwidth* for *k* modes is the smallest bandwidth producing a density with at most *k* modes, c.f. Silverman (1981). He thus used the Silverman critical bandwidth to test the club convergence hypothesis (2 modes) against the null hypothesis of global convergence (1 mode). *p*-values for the test were calculated through bootstrapping, resulting in support for Quah's twin peaks and vanishing middle-income group hypotheses (Bianchi (1997) Table I, p.402).

In Table 1, we replicated Bianchi (1997)'s results using our longer time series, thus finding support for the cross-sectional twin peaks hypothesis at the 5 % level of significance in 1980, 1990, and 1996. On the other hand, we cannot reject the null hypothesis of single-peakedness in 1961 and 1970, even at the 10% significance level. This suggests that the distributions in those different years were not drawn from the same (unique invariant) measure. This is not surprising in light of the stylized facts we highlighted in Section 2. However, it puts in question any testing methodology that focuses on multi-modality of cross-section distributions.

KOS (2001) formulated a more direct test of the convergence hypothesis by dividing the income distribution support into various income groups, and testing the null hypothesis of a single mode of the limiting distribution at one of those groups against the alternative of more than one. They failed to reject the null hypothesis of global convergence. However, the authors themselves pointed out that their test results are not robust to the (arbitrary) choice of income-groups. Moreover, their test is formulated for single vs. twin peaks, and would be difficult to extend to multiple peaks (e.g. a resilient middle-income group).

[15] In a third study, Paap and Dijk (1998) followed a different approach: They assumed the existence of two clubs, and focused on the selection of parametric densities for the rich and poor groups. Their choice of number of clubs to model is based on visual exploration of income distribution histograms with arbitrary bin-width selections. Consequently, there is very little overlap between their approach and ours.

A dynamic version of Bianchi's test of multi-modality can be developed along the lines of KOS (2001) to test for multi-modality of the limiting distribution of the estimated transition kernel. Using kernel-estimated transition densities, as described in Section 3, we can easily estimate limiting (invariant) densities for the transition kernel and test the null hypothesis of multiple modes based on these estimated invariance measure.

We first approximate our kernel-density estimated transition kernel by finite transition matrix $\tilde{P}_n$, and then approximate the fixed point of the former with the latter's fixed point $g_n^* = \tilde{P}_n g_n^*$.[16] We compute the latter simply by iterating on $\tilde{P}_n^s g$ for any initial transition density g, and $s = 1, 2, \ldots$, until convergence (in the sup-norm) over the fixed grid of $\boldsymbol{x} = (x_1, \ldots, x_n)$. We can then use Bianchi's bootstrapping approach by varying the bandwidth used in the original transition density estimation, defining critical bandwidth as the smallest possible value producing an invariant density with at most k modes.

The results using estimated 5-year and 10-year transitions are shown in Tables 2 and 3, respectively. For the 5-year transitions, we reject the null hypothesis of uni-modality (classical convergence) for the periods of 1966-71 and 1986-91 at the 5% significance level, but fail to reject the null in other periods. Consequently, the 5-year transition density results largely agree with the KOS results of global convergence. In contrast, conducting the test on 10-year transition densities, we reject global convergence for the period 1961-71, in favor of the convergence club hypothesis. Similarly, we seem to find four modes in the limiting distribution of the 10-year transition for the period 1971-81, and three modes for the period 1981-91. Consequently, we cannot draw any firm conclusions regarding the various convergence hypotheses based on this test. Taking into account Quah's (2001) critique of any statements about the limiting dynamics (convergence may take place after centuries of polarization), and taking into account the possibility of non-stationarity of the transition dynamics, we now turn to the issue of stochastic stability introduced in Sections 1-3.

---

[16] T. Li (1976) has shown in the case of the probability space $(X, \mathfrak{F}, \mu) = ([0,1], \mathfrak{F}, \lambda)$, where λ is Lebesgue measure, that each finite approximation P_n has a non-negative fixed point g_n^*, and $g_n^* \xrightarrow{n \uparrow \infty} g^*$ weakly. Bose (1994) expanded this result to strong convergence.

References

Azariadis, C. and Drazen, A., 1990. "Threshold externalities in economic development". *Quarterly Journal of Economics* 105, 501-526.

Barro, R. and Sala-i-Martin, X., 1992. "Convergence". *Journal of Political Economy* 100(2), 223-251

Barro, R. and Sala-i-Martin, X., 1995. *Economic Growth*, McGraw-Hill, Inc.

Bianchi, M., 1997. "Testing for convergence: evidence from nonparametric multimodality tests". *Journal of Applied Econometrics* 12(4), July-August, 393-409.

Bose, C., 1994. "On the existence and approximation of invariant densities for nonsingular transformations on $\mathbb{R}^d$". *Journal of Approximation Theory* 79, 260-270.

Bourguignon, F. and Morrison, C., 2002. "Inequality among world citizens: 1820-1992". *American Economic Review* 92(4), 727-744.

Bulli, S., 2001. "Distribution dynamics and cross-country convergence: a new approach". *Scottish Journal of Political Economy* 48, 226-243.

Desdoigts, A., 1999. "Patterns of economic development and the formation of clubs". *Journal of Economic Growth* 4(3), September, 305-330.

Durlauf, S., 1993. "Nonergodic economic growth". *Review of Economic Studies* 60, 349-366.

Durlauf, S. and Quah, D., 1999. "The new empirics of economic growth". In *Handbook of Macroeconomics*, (eds.) J. Taylor and M. Woodford, Vol. 1A, chapter 4., 231-304.

Fischer, S., 2003. "Globalization and its challenges". *American Economic Review* 93(2), 1-30.

Johnson, P., 2000. "A nonparametric analysis of income convergence across the US states". *Economics Letters* 69, 219-223.

Kremer, M., Onastski, A., and Stock, J., 2001. "Searching for prosperity". *Carnegie-Rochester Conference Series on Public Policy* 55, 275-303.

Li., T., 1976. "Finite approximation for the Frobenius-Perron operation: a solution to Ulam's conjecture". *Journal of Approximation Theory* 17, pp. 177-186.

Mankiw, N., Romer, D. and Weil, D., 1992. "A contribution to the empirics of economic growth". *Quarterly Journal of Economics*, 407-437.

Paap, R. and van Dijk, H., 1998. "Distribution and mobility of wealth of nations". *European Economic Review* 42(7), July, 1269-1293.

Phillips, P. and Sul, D., 2003. “The elusive empirical shadow of growth convergence”. *Cowles Foundation Discussion Paper* No.1398.

Quah, D., 1993. “Empirical cross-section dynamics in economic growth”, *European Economic Review* 37, 426-434.

_______, 1993. “Galton's Fallacy and tests of the convergence hypothesis”. *Scandinavian Journal of Economics* 95, 427-443.

_______, 1995. “Empirics for economic growth and convergence”. *European Economic Review* 40, 1353-1375.

_______, 1996a. “Convergence empirics across economies with (some) capital mobility”. *Journal of Economic Growth* 1, 95-124.

_______, 1996b. “Twin peaks: growth and convergence in models of distribution dynamics”. *Economic Journal* 106, 1045-1055.

_______, 1997. “Empirics for growth and distribution: polarization, stratification, and convergence club”. *Journal of Economic Growth* 2, 27-59.

_______, 2001. “Searching for prosperity: A comment”. *Carnegie-Rochester Conference Series on Public Policy* 55, 305-319.

Roussas, G., 1969a. “Nonparametric estimation in Markov processes”. *Annals of the Institute of Statistical Mathematics* 21, 73-87.

_________, 1969b. “Nonparametric estimation of the transition distribution function of a Markov process”. *Annals of Mathematical Statistics 40.* 1386-1400.

_________, 1991. “Estimation of transition distribution function and its quantiles in markov processes: strong consistency and asymptotic normality”. *Nonparametric functional estimation and related topics* (eds.) G. Roussas, 443-462.

Sala-i-Martin, X., 1995. “The classical approach to convergence analysis”. *Economic Journal* 106, 1019-1036

_____________ , 2002a. “The disturbing ‘rise’ of the global income inequality”. *NBER Working Paper* 8904.

_____________ , 2002b. “The world distribution of income (estimated from individual country distributions)”. *NBER Working Paper* 8933.

Silverman, B., 1981. “Using kernel density estimates to investigate multi-modality”. *Journal of Royal Statistical Society, Series B*, 43, 97-99.

< Table 1. Static multi-modality test >

	critical bandwidth				*p*-value			
	k=1	*k* =2	*k* =3	*k* =4	*k*=1	*k* =2	*k* =3	*k* =4
1961	0.3715	0.3560	0.2119	0.1784	0.380	0.053	0.287	0.154
1970	0.4020	0.3111	0.1827	0.1450	0.163	0.167	0.351	0.354
1980	0.4259	0.1637	0.1566	0.1283	0.041	0.867	0.541	0.461
1990	0.5572	0.2836	0.1468	0.0784	0.002	0.075	0.718	0.983
1996	0.5629	0.2864	0.2173	0.1482	0.000	0.090	0.077	0.208

Note: The *p*-value is based on bootstrapping with 1000 replications.

< Table 2. Dynamic multi-modality test, 5-year transition >

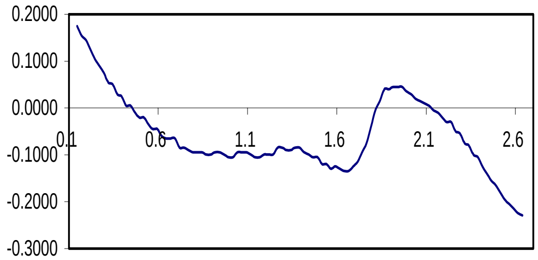
	critical bandwidth				*p*-value			
	k=1	*k* =2	*k* =3	*k* =4	*k*=1	*k* =2	*k* =3	*k* =4
1961-66	0.2803	0.1468	0.1326	0.1225	0.440	0.660	0.290	0.080
1966-71	0.4924	0.1386	0.1205	0.0978	0.040	0.750	0.400	0.290
1971-76	0.4338	0.1649	0.1249	0.0834	0.080	0.410	0.420	0.630
1976-81	0.4018	0.1948	0.1149	0.0895	0.180	0.270	0.420	0.480
1981-86	0.3837	0.0873	0.0720	0.0683	0.290	0.890	0.720	0.500
1986-91	0.5350	0.2040	0.1315	0.0747	0.040	0.390	0.330	0.750
1991-96	0.2355	0.0824	0.0804	0.0784	0.830	0.980	0.880	0.720

< Table 3. Dynamic multi-modality test, 10-year transition >

	critical bandwidth				*p*-value			
	k=1	*k* =2	*k* =3	*k* =4	*k*=1	*k* =2	*k* =3	*k* =4
1961-71	0.4652	0.2123	0.1874	0.1128	0.050	0.160	0.020	0.200
1971-81	0.4744	0.2796	0.2438	0.1365	0.040	0.050	0.000	0.020
1981-91	0.4440	0.2650	0.1149	0.0857	0.180	0.040	0.590	0.350

< Table 4. Test of three stable modes in 1961-66 >

	median	fixed value	s.e
C1	0.4500	0.4201	0.0890
C2	1.0185	1.4175	0.1357
C3	1.3078	1.6669	0.1768
C4	1.8115	1.8165	0.8501
C5	2.1357	2.0908	0.2334
Chi-squared Stat.			**12.9207**

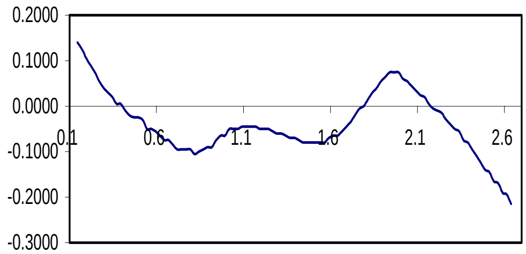


note: 1) s.e denotes standard error of median points.

2) Critical values of the Chi-square distribution with 5 degree of freedom: 10%=9.24, 5%=11.07, 1%=15.09

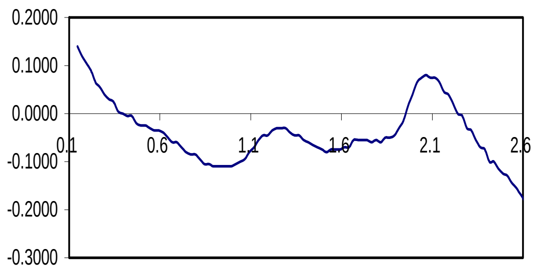
< Table 5. Test of three stable modes in 1966-71 >

	median	fixed value	s.e
C1	0.4001	0.3702	0.0756
C2	0.8290	1.0435	0.1097
C3	1.2130	1.6170	0.1393
C4	1.7916	1.7916	0.2676
C5	2.1706	2.1656	0.1806
Chi-squared Stat.			**12.3918**



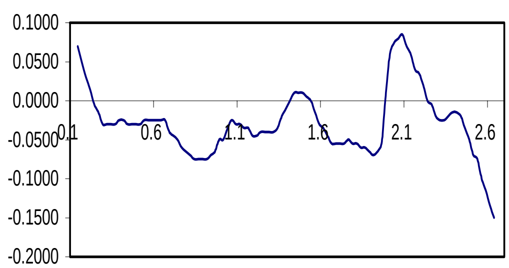
< Table 6. Test of three stable modes in 1971-76 >

	median	fixed value	s.e
C1	0.3951	0.3702	0.0872
C2	0.9737	1.1183	0.1450
C3	1.2779	1.4175	0.1669
C4	1.9262	1.9162	0.2999
C5	2.2404	2.2404	0.1633
Chi-squared Stat.			1.7768



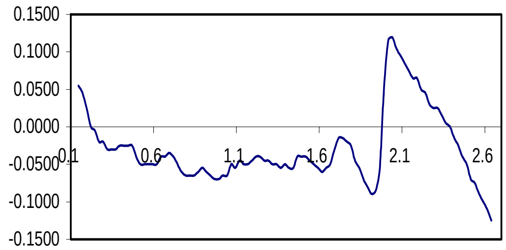
< Table 7. Test of three stable modes in 1976-81 >

	median	fixed value	s.e
C1	0.2405	0.2206	0.0417
C2	1.4175	1.2928	0.1329
C3	1.5422	1.4923	0.1026
C4	2.0010	1.9910	0.2708
C5	2.2404	2.2653	0.0972
Chi-squared Stat.			1.4117



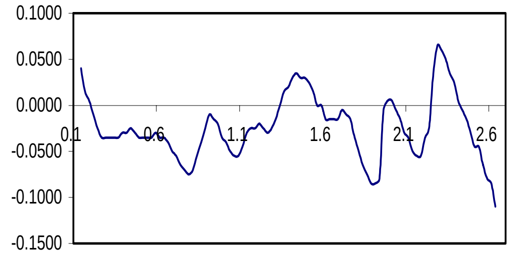
< Table 8. Test of three stable modes in 1981-86 >

	median	fixed value	s.e
C1	0.2206	0.2206	0.0378
C2	0.9786	1.2928	0.0904
C3	1.7866	1.6170	0.1361
C4	2.0409	1.9412	0.1848
C5	2.3900	2.3401	0.0843
Chi-squared Stat.			**14.2746**



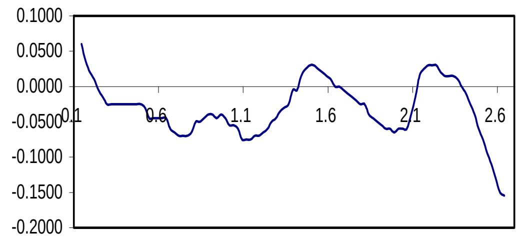
< Table 9. Test of three stable modes in 1986-91 >

	median	fixed value	s.e
C1	0.2106	0.1957	0.0302
C2	1.3028	1.2928	0.1303
C3	1.5921	1.5921	0.0901
C4	1.9960	2.0908	0.1297
C5	2.4349	2.4150	0.0731
Chi-squared Stat.			0.8577

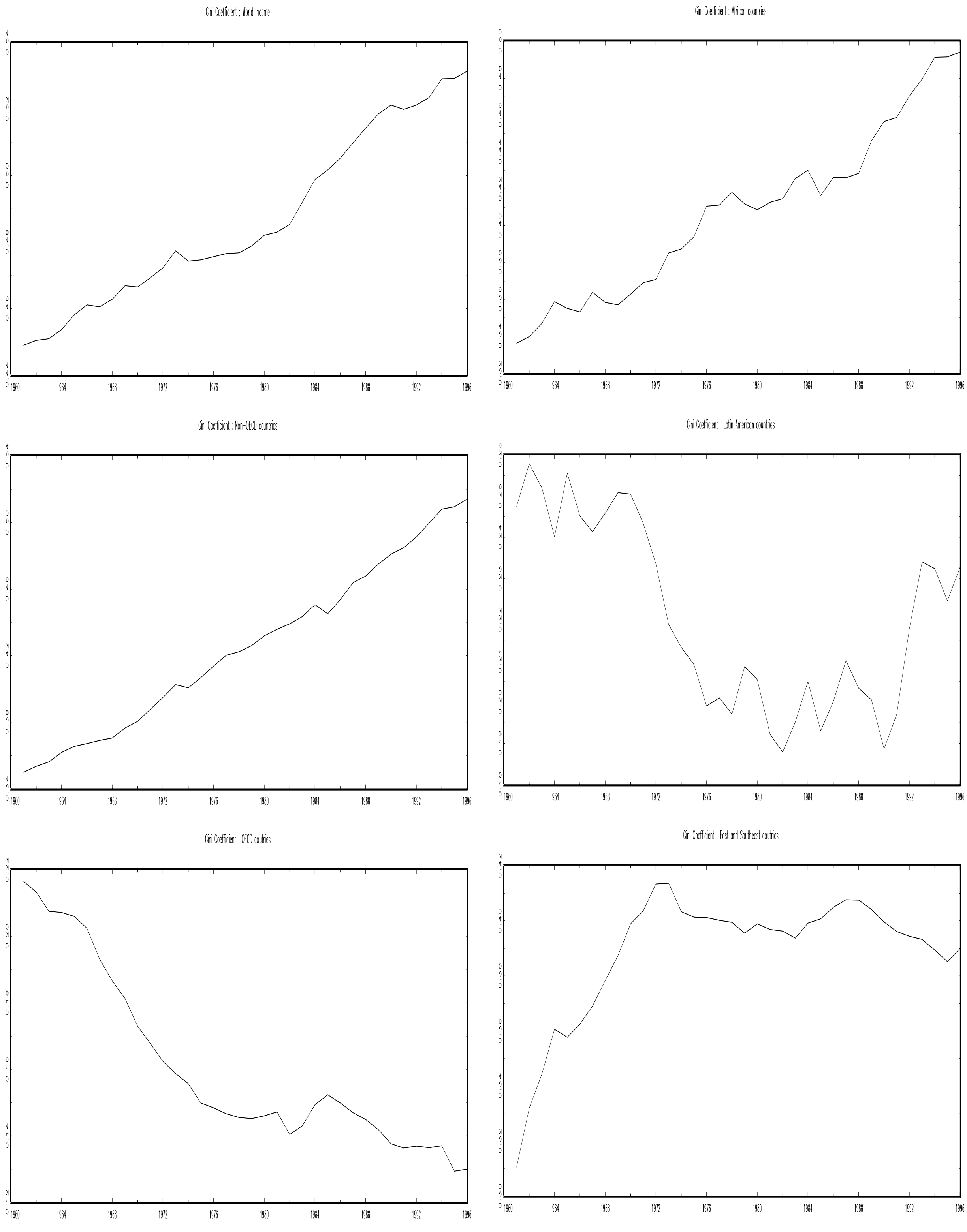


< Table 10. Test of three stable modes in 1991-96 >

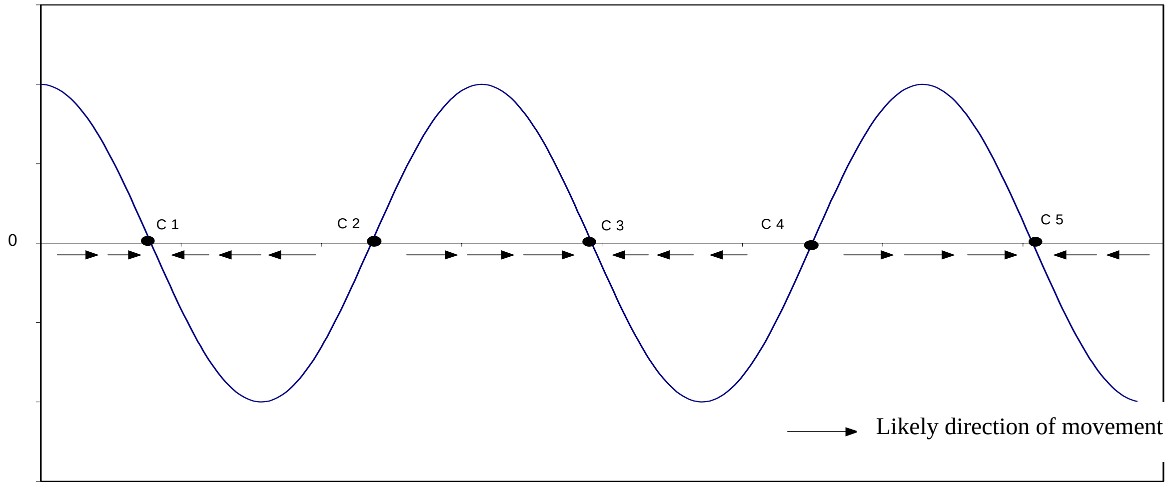
	median	fixed value	s.e
C1	0.2306	0.2206	0.0437
C2	1.4175	1.4175	0.1746
C3	1.6669	1.6918	0.1146
C4	2.1556	2.0908	0.2437
C5	2.3900	2.3651	0.1085
Chi-squared Stat.			0.2229



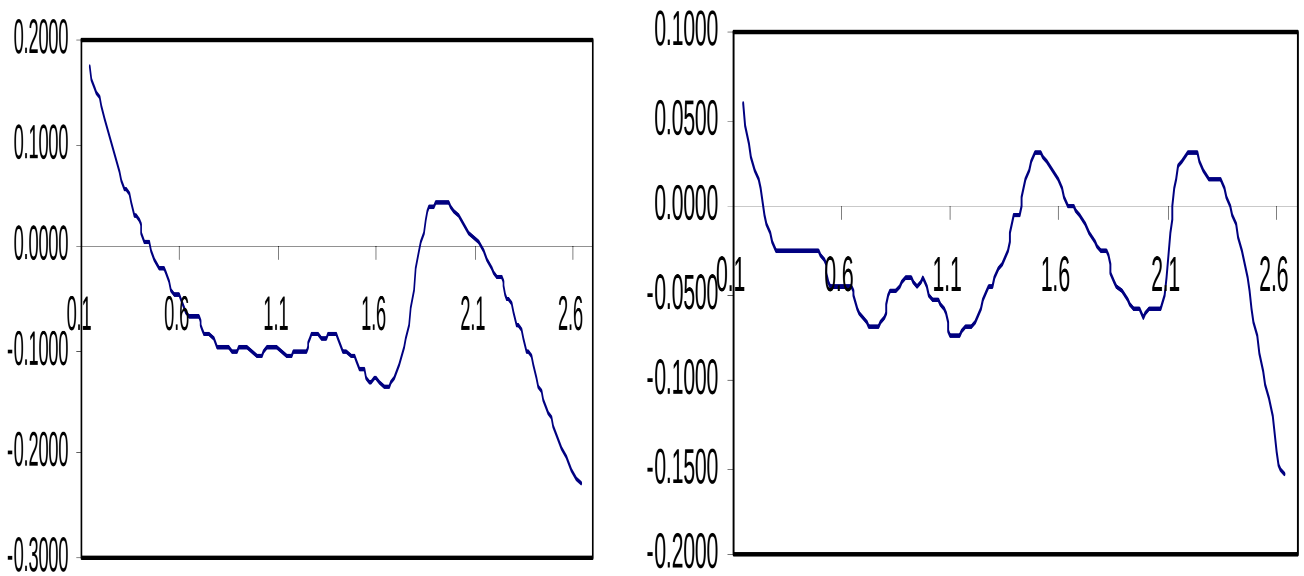
<Figure 1. Gini Coefficients: 1961-96 >



< Figure 2. Stochastic Stability 1>

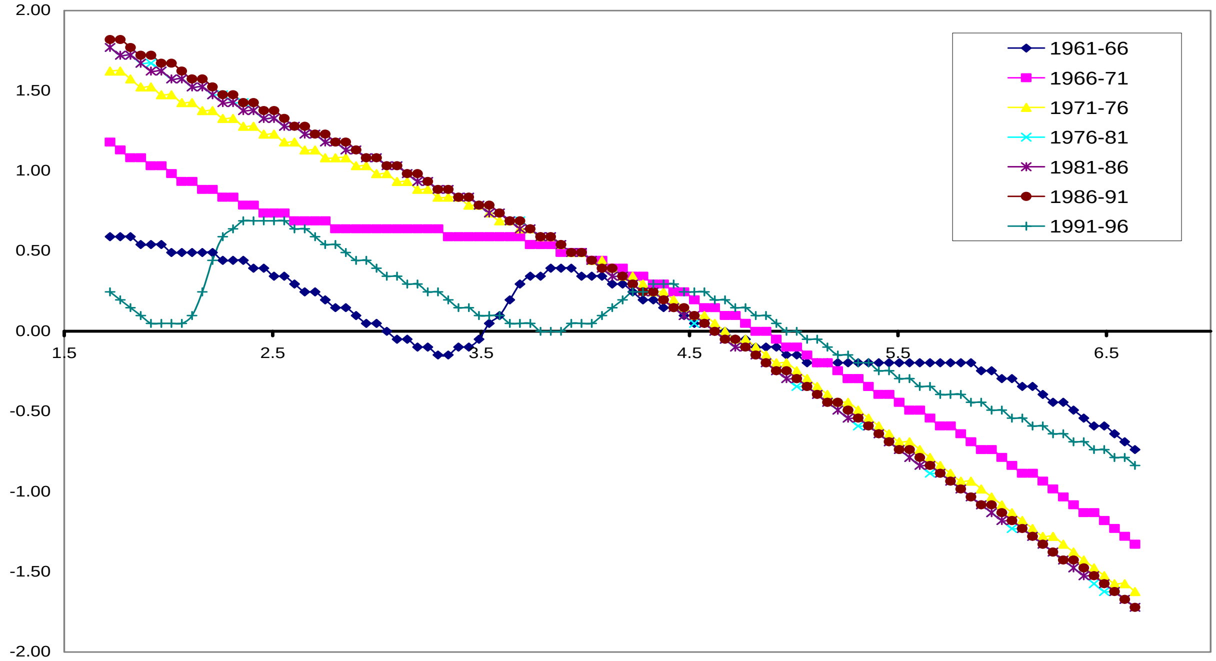


< Figure 3. Stochastic Stability2 >

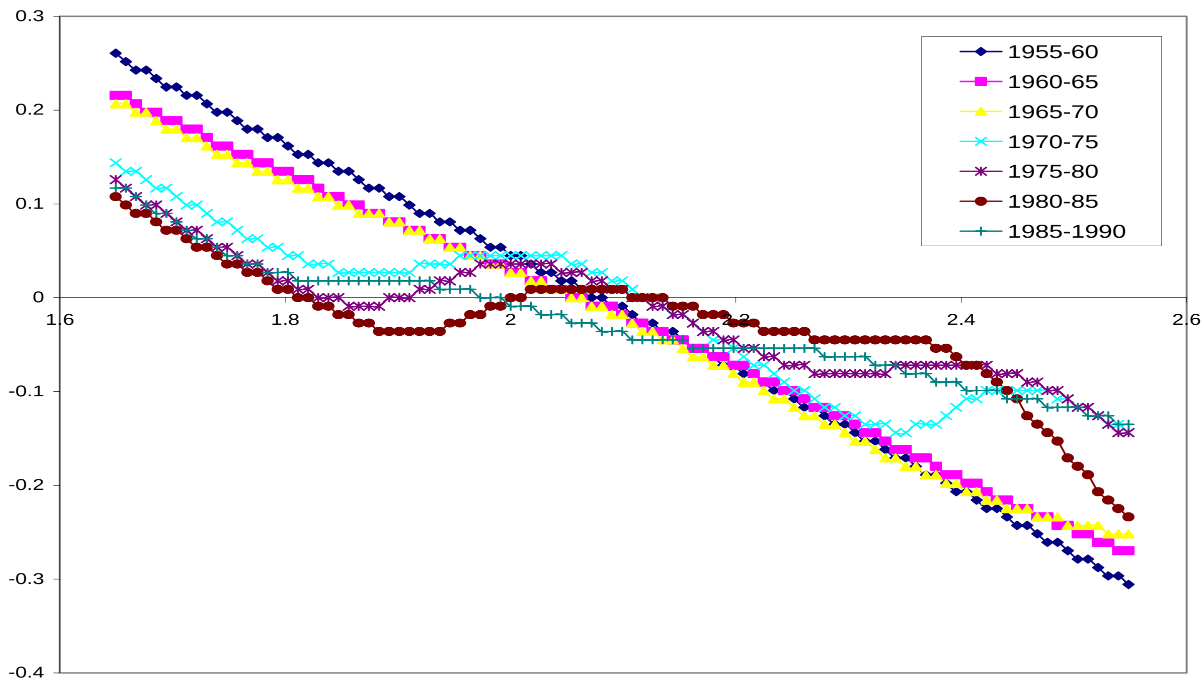


< Figure 4. A Single Stochastically Stable Zero >

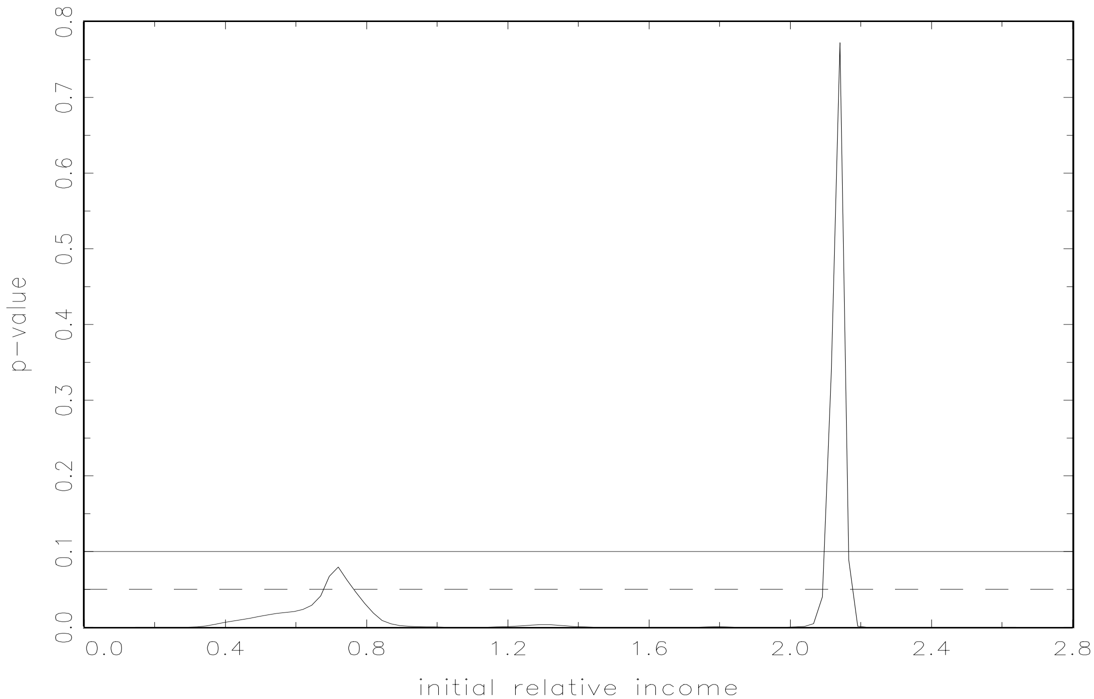
(a) OECD



(b) U.S. States



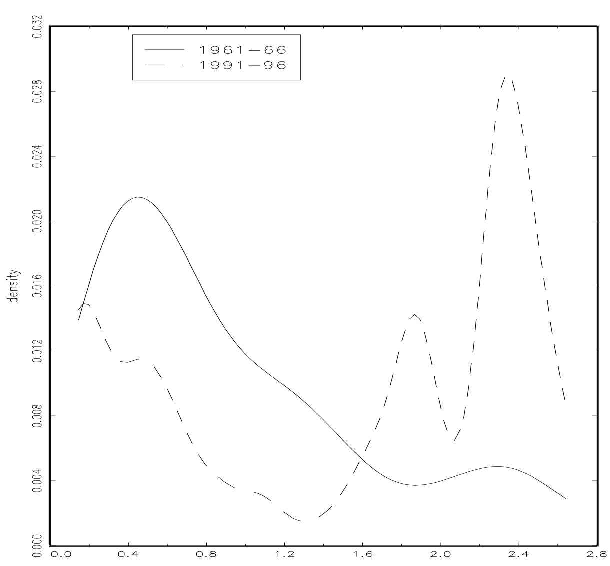
<Figure 5. *p*-values for Kolmogorv-Smirnov Tests of Equality of Transition CDF conditional on each relative income level: 1961-62 vs. 1995-96 >



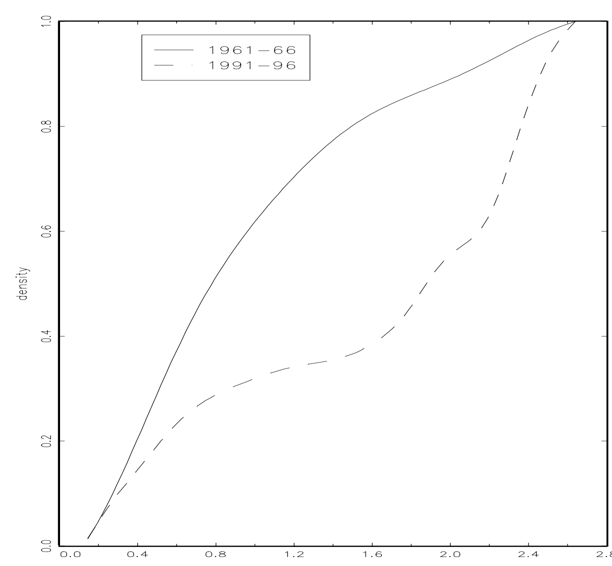
< Figure 6. Invariant (limiting) distributions for the transitions during 1961-66 and 1991-96>

p-value of Kolmogorov-Smirnov test of equality of limiting CDFs=1.3799889E-09

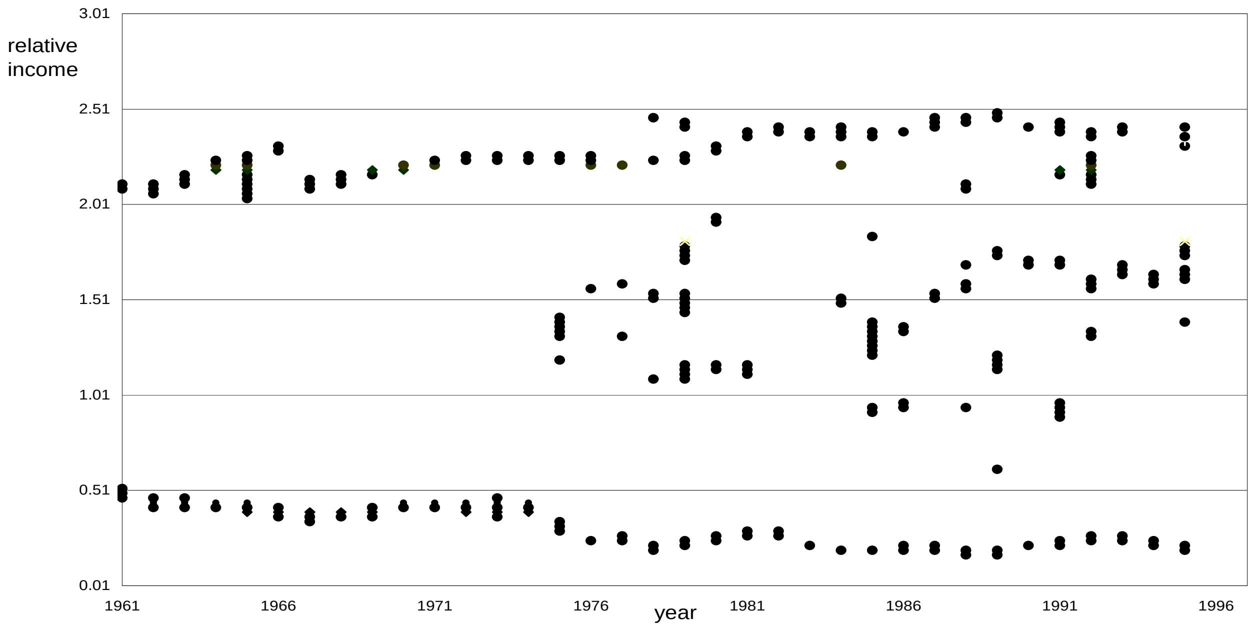
(a) Invariant (limit) pdfs



(b) Invariant (limit) CDFs



< Figure 7. Zeros of conditional median less previous period’s relative income >



< Figure 8. Probability of transition from Poor to Middle group in *n*-periods >