# Lessons from 40 years of cross-country convergence empirics

Paul A. Johnson*
Vassar College

Chris Papageorgiou
International Monetary Fund

Maria Grazia Pittau
Sapienza University of Rome

Roberto Zelli
Sapienza University of Rome

Prepared for the *Oxford Handbook of Income Distribution and Economic Growth*

January 2024

We survey the literature testing the absolute convergence hypothesis, the proposition that (the distributions of) countries' long-run per capita income levels are independent of their country-specific initial conditions. We conclude that the literature supports the view that the cross-country data is more consistent with the presence of several convergence clubs, groups of countries with similar initial conditions that tend to have similar long-run outcomes, than with absolute convergence, or a single convergence club. We revisit the data from 1970 to 2019 using a mixture model of the cross-country distribution of per capita income and find evidence of multiple convergence clubs. For the 2000s, and 2010s, this result is inconsistent with recent claims of convergence made by some researchers. We close with a consideration of future prospects for reductions in the gap in per capita incomes between poor countries and rich countries in light of the challenges posed by the Covid-19 pandemic, inflation and the associated financial tightening, climate change, and artificial intelligence.

JEL: E01, E13, O11, O47, F41, F62

We thank Marina Conesa and Hites Ahir for research assistance. Papageorgiou acknowledges financial support from the U.K.'s Foreign, Commonwealth & Development Office (FCDO). The views expressed in this paper are those of the authors and do not necessarily represent those of the International Monetary Fund, IMF policy or FCDO.

*Correspondence to: Paul A. Johnson, Department of Economics, Vassar College, 124 Raymond Avenue, Poughkeepsie, NY 12604, USA (pajohnson@vassar.edu)

## 1. Introduction

Interest in convergence in the modern growth literature begins with interest in catching-up behavior based on the relationship between initial income and subsequent growth.[1] According to this notion, two countries exhibit convergence if the one with lower initial income subsequently grows faster than the other and so tends to "catch up" to the higher income country. One mechanism driving catching-up is the possibility of rapid technological progress as countries behind the technological frontier can move towards it simply through investment in physical capital embodying frontier technologies rather than having to discover anew such technologies for themselves. The earliest modern studies of catch-up convergence such as Abramovitz (1986), Baumol (1986) and Marris (1982) emphasized this mechanism. However, following the work of Barro (1991), Barro and Sala-iMartin (1992) and Mankiw, Romer and Weil (1992), among others, catch-up convergence became known as $\beta$-convergence in reference to the key parameter in the oft-estimated growth regression and the role of diminishing returns in long-run behavior in the neoclassical growth model became the more prominent mechanism driving catching-up.[2] So much so that Mankiw (1995, p. 301) argues that for "understanding international experience, the best assumption may be that all countries have access to the same pool of knowledge, but differ by the degree to which they take advantage of this knowledge by investing in physical and human capital." By contrast, Dowrick and Rogers (2002) argue that both diminishing returns and technology transfer are important contributors to the convergence process.[3]

In addition to the long-standing intrinsic interest in international inequality among economists and others, another motivation for interest in convergence was the belief that its presence or absence could adjudicate between the class of neoclassical models that

---

[1]See Durlauf et. al. (2009) for references to earlier interest in the convergence or divergence of nations.
[2]Motivated by the neoclassical model, Kormendi and Meguire (1985) include initial income levels in their growth regressions and note the catching-up implications of the negative estimated coefficient despite not being primarily concerned with the convergence hypothesis per se.
[3]See also Bernard and Jones (1996) and Barro and Sala-i-Martin (1997).

feature decreasing returns to capital and exogenous technological progress, and predict catching up by low income countries and the class of endogenous growth models typified by Romer (1986) and Lucas (1988) which predict that high initial income countries will tend to grow faster than low initial income countries. However, not all endogenous growth models imply the absence of catching up behavior. Jones and Manuelli (1990) and Kelly (1992), for example, present endogenous growth models compatible with catching up. Moreover, as we discuss below, $\beta$-convergence tests can suggest the presence of catching up even when convergence is not occurring because of the low power of those tests in some circumstances as discussed by Bernard and Durlauf (1996). All things considered, the belief that $\beta$-convergence tests could reveal whether the growth process was better described by neoclassical or endogenous growth models proved to be mistaken.

Fundamental to the many empirical investigations of convergence in the cross-country context has been the availability of internationally comparable data on the relevant variables for a broad cross-section of countries, initially due to the work of Kravis et. al. (1978) and Summers and Heston (1988, 1991) and now available as the Penn World Table (Feenstra et. al., 2015). Much of the literature employs one version of this data set or another.[4]

This paper surveys the empirical literature on cross-country convergence from the past 40 years with a view to identifying the lessons that can be drawn from that literature. Our focus is the absolute convergence hypothesis, the proposition that (the distributions of) countries' long-run per capita income levels are independent of their country-specific initial conditions. Throughout we draw extensively on prior work including Durlauf et. al. (2005) and Johnson and Papageorgiou (2020).[5] We argue that the literature

[4]For an analysis of the sensitivity of empirical results to the PWT version employed see Johnson et. al. (2013).

[5]Other surveys of aspects of the empirical growth literature include Durlauf and Quah (1999), Temple (1999), and Islam (2003).

overwhelmingly supports the view that convergence, in the sense of long-run per capita income levels being independent of initial conditions, is not occurring. Instead, the wide variety of tests that have a long-run role for initial conditions in the determination of per capita incomes as an alternative hypothesis typically reject the null hypothesis of convergence.[6]

Notably, we find recent claims of short-term $\beta$-convergence during the 20 years or so prior to the COVID-19 pandemic to be unconvincing objections to the view that long-run per capita income levels are dependent on initial conditions. Rather, we argue that that view is well supported by data from the 2000s and 2010s. Moreover, we do not expect whatever small amount of catching up that may have occurred in the 2000s and 2010s to continue into the 2020s and beyond as developing countries are likely to fair far less well that developed countries in the face of the currently visible set of headwinds and opportunities.

The plan of the paper is as follows: Section 2 provides a definition of convergence and discusses many of the statistical approaches to testing the hypothesis that have been used before taking a preliminary look at the cross-country data for the 1970 to 2019 period; Section 3 surveys the literature that has employed the statistical approaches introduced in Section 2; Section 4 revisits the data with a view to underscoring our view that the literature supports the view that countries' long-run per capita income levels are not independent of their country-specific initial conditions i.e. that convergence is not occurring; Section 5 considers future prospects for reductions in the gap in per capita incomes between poor countries and rich countries in light of the challenges posed by the COVID-19 pandemic, inflation and the associated financial tightening, climate change, and artificial intelligence; and, Section 6 offers some concluding remarks.

[6] As we emphasize later, strictly speaking, it is not possible to unequivocally distinguish between the two views with a finite span of data. See Durlauf et. al. (2005, pp 622-3).

## 2. Preliminaries

### 2.1 Convergence concepts and convergence tests

The convergence hypothesis states that a country's initial conditions have no effect on (the distribution of) its long-run per capita income level. The primary theoretical motivation for this proposition is the standard neoclassical growth model which, under suitable regularity conditions, has a unique, stable steady state. This means that, provided a country starts with a non-zero quantity of capital per worker, output per worker will converge to steady-state level independent of initial capital per worker.[7] The mechanism driving convergence is the assumed diminishing returns to capital, or more generally, to all accumulable factors of production. Applied to a group of countries, the long-run irrelevance of initial conditions is usually taken to mean that the countries' per capita incomes ought to be getting closer to each other in some sense which implies that poorer countries are catching up with richer countries. In this subsection we give a somewhat more precise definition of convergence before outlining some of the econometric approaches to testing the hypothesis. The following subsection takes a preliminary look at the behavior of the cross-country data on per capita income since 1970.

Letting $y_{i,t}$ denote per capita income in country $i$ in period $t$ and $\Omega_t$ denote history until time $t$ (i.e. initial conditions going forward), one way to express the irrelevance of initial conditions for long-run differences per capita income between countries $i$ and $j$ is to require that

$$\lim_{\tau \to \infty} E(\log y_{i,t+\tau} - \log y_{j,t+\tau} | \Omega_t) = 0 \tag{1}$$

[7]See Barro and Sala-i-Martin (2003), Acemoglu (2009), Aghion and Howitt. (2009), Weil (2014), or Romer (2018), among others, for an exposition of the neoclassical growth model.

for any $\Omega_t$ with $y_{i,s} \in \Omega_t$ for all $i$ and for all $s \leq t$, meaning that any differences in the per capita incomes of countries $i$ and $j$ are expected to eventually disappear.[8]

More generally, convergence can be defined in terms of the entire distribution of per capita income as

$$\lim_{\tau \to \infty} \| F_{i,t+\tau}(x|\Omega_t) - F_{j,t+\tau}(x|\Omega_t) \| = 0 \tag{2}$$

where $F_{i,t+\tau}(x|\Omega_t)$ is the distribution of per capita income in country $i$ at time $t+\tau$ conditional on $\Omega_t$ and $\| \ \|$ denotes a metric for computing the distance between distributions.[9]

A stochastic version of the neoclassical growth model would, under suitable regularity conditions, have a unique, stable stochastic steady state and so satisfy both Equation (1) and Equation (2) as, if all countries were to satisfy such a model, they would all have the same stochastic steady state. Equation (1) above states that any differences in the per capita incomes of countries $i$ and $j$ are expected to eventually disappear so, for $y_{i,t} > y_{j,t}$, it implies that $E(\log y_{i,t+\tau} - \log y_{j,t+\tau} | \Omega_t) < \log y_{i,t} - \log y_{j,t}$ for some $\tau > 0$, so that

$$E\left[ \frac{\log y_{i,t+\tau} - \log y_{i,t}}{\tau} - \frac{\log y_{j,t+\tau} - \log y_{j,t}}{\tau} \Big| \Omega_t \right] < 0 \tag{3}$$

for that $\tau$. That is, the country with higher initial income will have a lower average growth rate over some future time interval if convergence is occurring. That is, Equation

[8]This definition, suggested by Bernard and Durlauf (1995, 1996), does suffer from the weakness that it does not control for long-run deviations whose current direction is not predictable as would be the case were the $\log y_{i,t} - \log y_{j,t}$ process a random walk with time $t$ value of zero so that $\log y_{i,t+\tau} - \log y_{j,t+\tau}$ could become arbitrarily large for some $\tau$. See Hall, Robertson and Wickens (1997) for more on this issue which we ignore.

[9]Durlauf et. al. (2005) exposit various definitions of convergence, the relationships between them, and their empirical implications. Durlauf et. al. (2005, 2009) provide extensive discussions of the econometric issues that arise in convergence testing.

(1) implies "catching-up", the motivation for $\beta$-convergence tests of the convergence – tests that check for a negative correlation between $\frac{\log y_{i,t+\tau} - \log y_{i,t}}{\tau}$ and $\log y_{i,t}$ in a cross section of countries over some time period.

This catch-up notion of convergence is often operationalized by using data on a cross-section of countries to estimate an equation such as

$$g_{i,t,\tau} = \alpha + \beta \log(y_{i,t}) + \epsilon_i \tag{4}$$

where $g_{i,t,\tau} = 100(\log y_{i,t+\tau} - \log y_{i,t})/\tau$ is the average growth rate of per capita GDP in country $i$ over the $\tau$ years subsequent to year $t$. A value of $\beta < 0$ would imply that, on average, poorer countries grew faster than richer countries between $t$ and $t + \tau$ so the catch-up operationalization of the convergence hypothesis is investigated by testing the null hypothesis $\beta = 0$ against the alternative $\beta < 0$ with rejection being considered evidence of catch-up or so called $\beta$-convergence.

The converse of the proposition that Equation (1) implies Equation (3) is not true, however, as the reduction of the gap in per capita income levels over some interval carries with it no implication of long-run equality of expected per capita income levels. This inability of $\beta$-convergence tests to reliably refute the convergence hypothesis when Equation (1) does not hold reflects, of course, the well-known low power of $\beta$-convergence tests against at least some non-convergent alternatives as noted by Bernard and Durlauf (1996) and Durlauf and Johnson (1995) among others.

As Durlauf et. al. (2005) show, it is possible to offer a versions of the definitions above that include control variables permitting some types of cross-country heterogeneity. Some researchers (e.g. Barro, 1991, and Mankiw et. al., 1992) estimate versions of Equation (4) that include controls such as different rates of physical and human capital accumulation. Doing so changes the convergence concept from "absolute" $\beta$-convergence to "conditional" $\beta$-convergence. Conditional $\beta$-convergence implies that a poor country can catch up simply by adopting the appropriate characteristics of a rich country. Dowrick and DeLong (2003, p. 204) describe the presumption that a poor country could do this as the "joker in the deck," arguing that "a moment's thought will convince anyone that many of the right-hand-side variables used by Barro (1997) could never be brought to the mean values found in the industrial core of the world economy in any country that has not already attained the productivity level and socioeconomic structure found in the industrial core." Accordingly, we consider only absolute or unconditional approaches to convergence in this paper. Quah (1996) also argues that it is unconditional rather than conditional convergence that ought to be the hypothesis of interest.[10]

An important class of alternatives to convergence arises in growth models that have multiple stable steady states which can occur for a variety of reasons such as the presence of a feedback loop from per capita income to some other state variable.[11] These models do not exhibit convergence as defined in Equation (1) as an economy obeying such a model will converge to the steady state associated with the basin of attraction in which it begins. This means that its long-run per capita income will depend on its initial conditions as measured by the state variables indicating basin of attraction membership. Economies having initial conditions that are sufficiently similar for them to lie in the same basin of attraction will converge locally and so form a "convergence club". The existence of two or more such clubs means that global differences in per capita incomes will persist indefinitely and, taken together, a group of economies obeying such a model

---

[10]We are also are mindful of the endogeneity issues brought by such control variables as discussed by Cho (1996), Temple (1999), Easterly (2005), Durlauf, Johnson, and Temple (2005), Rodrik (2012), and Lenkoski, Eicher, and Raftery (2014), among others.

[11]Azariadis (1996), Galor (1996), and Johnson and Papageorgiou (2020) provide surveys of theoretical mechanisms capable of producing multiple stable steady states. These include different saving propensities out of labor and capital income; low elasticities of substitution between capital and labor; demographic transitions to sharply lower fertility rates as wages rise; external increasing returns; and, external effects from social interactions, to list a few. Especially relevant to our later analysis is Bloom, Canning and Sevilla (2003) who derive a mixture reduced form from an explicit economic model with multiple steady states.

will exhibit so called club-convergence with club membership depending on initial conditions.[12]

Suppose, for example, that there are two basins of attraction, *Rich* and *Poor* with $\lim_{\tau\to\infty} E(\log y_{i,t+\tau}|i \in \mathit{Rich}) = \mu_R$ and $\lim_{\tau\to\infty} E(\log y_{i,t+\tau}|i \in \mathit{Poor}) = \mu_P < \mu_R$ where $i \in \mathit{Rich}$ indicates that country $i$ falls in the *Rich* basin of attraction and $i \in \mathit{Poor}$ indicates that country $i$ falls in the *Poor* basin of attraction, then $\lim_{\tau\to\infty} E(\log y_{i,t+\tau} - \log y_{j,t+\tau}|\Omega_t) \neq 0$, violating Definition (1) whenever $i \in \mathit{Rich}$ and $j \in \mathit{Poor}$ or vice versa. Assuming, for ease of exposition, a common value of $\beta$, the version of Equation (4) that results would have two different values of $\alpha$: $\alpha_P$ for countries in the *Poor* basin of attraction, and $\alpha_R$ for the countries *Rich* basin of attraction with $\alpha_P < \alpha_R$, so that,

$$g_{i,t,\tau} = \begin{cases} \alpha_P + \beta \log(y_{i,t}) + \epsilon_i & \text{if } i \in \mathit{Poor} \\ \alpha_R + \beta \log(y_{i,t}) + \epsilon_i & \text{if } i \in \mathit{Rich} \end{cases}.$$

Bernard and Durlauf (1996) show that, even ignoring sampling error, there is no guarantee that the estimated value of $\beta$ will not be negative if Equation (4) is estimated in this case which can cause erroneous conclusions in favor of convergence despite the model's violation of Equation (1). In general, $\beta$-convergence tests have low power against the club convergence alternative.

Another statistical implementation of the proposition that initial conditions play no role in long-run outcomes, with its implication that contemporary differences in per capita incomes are transitory, is based on the claim that the dispersion of per capita incomes across economies should fall if catching up is occurring. Letting $\sigma_t^2$ denote the cross-country variance of $y_{it}$ (or, in some applications, of $\log y_{it}$), Barro and Sala-i-Martin

[12]Another way to think of the issue of the low power of $\beta$-convergence tests is that initial income need not be a sufficient statistic for a country's initial conditions.

(1992) refer to this concept as $\sigma$-convergence which is said to occur between $t$ and $t+\tau$ if $\sigma_t^2 > \sigma_{t+\tau}^2$. Young et. al. (2008) show that $\beta$-convergence is necessary for $\sigma$-convergence, however, as shocks can cause to $\sigma_t^2$ be constant or increase over time even if $\beta$-convergence is occurring, $\beta$-convergence is not sufficient for $\sigma$-convergence.[13] Believing otherwise is known as Galton's fallacy, as Friedman (1992), Quah (1993b), and Hart (1995) point out.

As with $\beta$-convergence, a finding of $\sigma$-convergence, can occur even if convergence in the sense of the irrelevance of initial conditions is not occurring. It is easy to construct examples with multiple stable steady states, so that Equation (1) is not satisfied, and yet the variance of per capita GDP across all countries falls over long periods of time as countries converge locally within their respective basins of attraction. That is, in general, $\sigma$-convergence tests that simply seek a declining $\sigma_t^2$ will also have low power against the alternative hypothesis of convergence clubs.

Phillips and Sul (2007a, 2007b, 2009) formalize the analysis of the dynamic behavior of the cross-country dispersion in per capita incomes and provide a test of "relative convergence", which is closely related to $\sigma$-convergence tests. They hypothesize a steady-state growth path common to all countries denoted as $\mu_t$ and write log $y_{it} = b_{it}\mu_t$ where $b_{it}$ describes the transition path of economy $i$ to the steady state growth path. Defining $h_{it} = n\log y_{it} \Big/ \sum_{j=1}^{n} \log y_{jt} = nb_{it} \Big/ \sum_{j=1}^{n} b_{jt}$ where $n$ is the number of countries, convergence is said to occur if $\lim_{t\to\infty} h_{it} = 1$ for all $i$. A measure of the cross-section sample variation of $h_{it}$ is then calculated as $H_t = \frac{1}{n}\sum_{i=1}^{n}(h_{it}-1)^2$ and Phillips and Sul propose the "log $t$" convergence test, a conventional one-sided t-test of the hypothesis $\gamma = 0$ (no convergence) against the alternative $\gamma > 0$ (convergence) in the regression

[13]See, for example, Lichtenberg (1994) for discussion of this point. See also Furceri (2005).

$$\log \frac{H_1}{H_t} - 2\log(\log t) = a + \gamma \log t + u_t.$$

Phillips and Sul (2007) show that $\gamma < 0$ implies divergence, $0 < \gamma < 2$ implies convergence in growth rates, and $2 \leq \gamma$ implies convergence in levels.

Kong et. al. (2019 and 2020) develop a related "weak $\sigma$-convergence" test based on the linear regression of the sample cross-section variance of $\log y_{it}$ on a time trend. As these authors observe, among other issues, such a test addresses the issue that the decline in cross-section variation of the variable of interest need not be monotonic because of shocks. The exact relationship between weak $\sigma$-convergence and relative convergence depends on the trend behavior of the common factor in the $\log y_{it}$ with weak $\sigma$-convergence generally being more generally applicable despite being more restrictive in some circumstances.[14]

Other approaches also seek to exploit the time series variation in the cross-country data. For example, Bernard and Durlauf (1995, 1996) observe that one implication of Equation (1) is that, if $\log y_{i,t}$ and $\log y_{j,t}$ obey integrated processes, then they will cointegrated with cointegrating vector $[1, -1]$.[15] Pesaran (2007) is one of many subsequent researchers who studies convergence as cointegration. Bernard and Durlauf (1996) argue that the cross-section approaches to convergence, such as $\beta$-convergence, are more appropriate than the time-series tests when economies are in transition to their steady states while the time series tests are more appropriate when economies are in their stochastic steady states. Papageorgiou and Perez-Sebastian (2004) argue that transition dynamics are an important source of variation in the cross-country per capita income data.

The understanding of the economic importance of the alternative hypothesis of club convergence and the desire for tests that have power against that alternative has given rise to a variety of statistical approaches that admit as an alternative the sort of

[14]Sul (2019) presents a comparison of both relative convergence and weak $\sigma$-convergence.
[15]Here we are ignoring the pathological case discussed in footnote 4.

clustering of countries indicative of convergence clubs. An early example is Durlauf and Johnson (1995) who use regression trees to sort the countries in the Mankiw et. al. (1992) dataset into groups of countries that are locally convergent based on initial conditions as measured by initial per capita income and literacy levels.[16] This yields a test of the convergence hypothesis that has more power against the club convergence alternative than the conventional (conditional) $\beta$-convergence tests in Mankiw et. al. (1992). Subsequent work that allows for the possibility that data-selected groups of countries will exhibit different long-run behavior includes Desdoigts (1999) and Kourtellos (2002) who use projection pursuit methods, Canova (2004) who takes a predictive density approach, Owen et. al. (2009) and Di Vaio and Enflo (2011) who use mixture models, and Jerzmanowski (2006) and Kerekes (2012) who use Markov switching models.

Noting the economic importance of the alternative hypothesis of club convergence, Phillips and Sul (2007b, 2009) provide an algorithm for sorting the data into potential locally convergent groups and testing the hypothesis of convergence against the club convergence alternative.[17] Other time-series approaches to data selection of groups of countries potentially indicative of convergence clubs include Hobijn and Frances (2000) and Beylunioglu et. al. (2020) who extend Pesaran's (2007) pair-wise approach using a "maximal clique" method. One difficulty with these approaches is that the data include countries which have converged as well as countries that are converging and, as Bernard and Durlauf (1996) suggest, time series approaches to testing the convergence hypothesis may be more applicable to the former than to the latter.

---

[16]Other early examples of sample splitting approaches include Baumol and Wolff (1988), Grier and Tullock (1989), and Dowrick (1992). The earliest research using data-dependent splitting of which we are aware is Chatterji (1992) who estimates a version of Equation (4) that includes the square and the cube of $\log y_{i,t}$ as explanatory variables and concludes that the implied nonlinear difference equation for $y_{i,t}$ (relative to the US value) has two basins of attraction with countries belonging to one or the other depending on whether their initial income per capita is greater or less than about one-sixth of the US value.
[17]Tomal (2023) provides a useful survey of the theory and applications of the clustering algorithm described in Phillips and Sul (2007, 2009). Kong et. al. (2019) briefly discuss the need for a method of identifying convergence clubs in the context of weak $\sigma$convergence but leave the design and implementation of a suitable algorithm for future research.

In a highly influential series of papers, Quah (1993, 1996a, 1996b, 1996c, 1997, 1999) criticizes the $\beta$-convergence and $\sigma$-convergence approaches to testing the convergence hypothesis as being unsuitable for the study of key issues such as mobility, stratification, and polarization in the cross-country distribution of per capita output. As an alternative, Quah proposes studying the entire cross-country distribution of per capita output and its evolution over time rather than just the behavior of representative country as in $\beta$-convergence tests or a single moment of the distribution as in $\sigma$-convergence tests.[18] Quah's emphasis on the entire cross-country distribution of per capita output and his stylized fact that the long-run distribution exhibits "twin peaks" – two modes – indicative of countries accumulating at low and high levels of income created the connection between convergence clubs and multimodality in the cross-country distribution of per capita income.

Convergence clubs can result in multiple modes in the cross-country distribution of per capita income as the countries in each convergence club move closer to each other resulting in accumulations of mass. Multimodality is, however, not sufficient to imply that convergence clubs exist as existence of convergence clubs implies limits on the mobility within the distribution as countries remain in their respective basins of attraction and the shape of the distribution *per se* does not restrict mobility within it. Quah (1996, 1997), Bianchi (1997), Jones (1997), Henderson et al. (2008), Krause (2017), and others have used kernel estimation methods to study the shape of the cross-country distribution of per capita income. Bianchi (1997) and Henderson et al. (2008) present tests of the hypothesis of a unimodal distribution against that of a multimodal distribution. Rejecting the null and finding little mobility between the modes they identify, their results are consistent with the existence of convergence clubs.

---

[18]Quah labeled the study of the time-series evolution of distributions "distribution dynamics", a field that has had many contributions including Johnson (2000), Fiaschi and Lavezzi (2003a, 2003b and 2007a, 2007b), Fotopoulos (2006, 2008), Maasoumi, Racine, and Stengos (2007), Fischer and Stumpner (2008), and Bandyopadhyay (2004, 2011).

One problem with this approach is that there is no reason why multiple basins of attraction will manifest as multiple modes in the cross-country distribution of per capita income. An improvement is to model the distribution as a finite mixture of other distributions i.e., as the weighted average of a finite number of component densities with specified functional form, usually the normal. Describing the cross-country distribution of per capita income in this way, allows the components to be interpreted as corresponding to the basins of attraction in the dynamic process describing the evolution of per capita income so that multiple components can be indicative of multiple basins of attraction. As the number of component densities can be data-determined, the mixture approach can thus provide a convergence test with more power against the club convergence alternative than the kernel approach as it is able to detect the presence of multiple components in a distribution even if that multiplicity does not manifest itself as multimodality.[19]

Of course, as with the detection of multiple modes in the kernel density approach, if multiple components are detected, their interpretation as indicative of convergence clubs requires an analysis of the mobility within the distribution, which in this case means that between the components. The mixture model framework provides a natural way to address this issue as the estimated parameters enable computation of the conditional probability that a country belongs to each component. Countries can thus be assigned to the component to which that have the highest probability of belonging and the propensity of countries to change their assigned components over time provides a measure of within-distribution mobility.

Paap and van Dijk (1998) estimate a mixture model of the cross country distribution of per capita income fixing the number of components at two based on the bimodality of histograms of their data. Later work such as Pittau et. al. (2010) and

[19]The parameters to be estimated in a mixture model of a probability distribution are the number of, the weights attached to, and the parameters of, the component densities.

Vollmer et. al. (2013) find evidence of three components when the number of components is determined in the course of the model estimation.[20]

Much has been learned about convergence or lack thereof in the cross-country distribution of per capita income through applications of the methods described above. In the following subsection we take a preliminary look at our dataset using some of these methods before returning to the literature with a view to drawing conclusions about the veracity of the convergence hypothesis.

### 2.2 A first look at the data

Here we take a first look at the evolution of the cross-country distribution of per capita income during the 50 years prior to the Covid-19 pandemic. Our data is a balanced panel of 103 countries over the 1970-2019 period from PWT 10.01 (Feenstra et. al. 2015).[21] Using the PWT mnemonics, we measure per capita income as *RGDPO*/*POP*, where *RGDPO is* output-side real GDP at chained PPPs in millions of 2017 USD and *POP* is the population in millions. To construct the balanced panel we exclude all countries without data on *RGDPO* and/or *POP* for the entire period, and, following standard practice in the literature, the middle-eastern oil producing countries, Luxemburg, and countries with populations of less than one million persons.[22,23] The data in each year are normalized by the sample average of per capita income for that year computed

---

[20]See also Alfó et. al. (2008), Battisti and Parmeter (2013), Pittau et. al. (2016), and Anderson et. al. (2016).

[21]Starting our data in 1970 gives a good balance between the cross-section and time dimensions of the dataset. Starting in 1960 would mean losing 16 countries to gain a decade's data, while starting in 1980 would mean having just one more than the 1970 start at the cost of losing a decade's data.

[22]The PWT mnemonics of the included countries are ALB, ARG, AUS, AUT, BDI, BEL, BEN, BFA, BGD, BGR, BOL, BRA, CAF, CAN, CHE, CHL, CHN, CIV, CMR, COD, COL, CRI, DEU, DNK, DOM, ECU, EGY, ESP, ETH, FIN, FRA, GBR, GHA, GIN, GRC, GTM, HKG, HND, HTI, HUN, IDN, IND, IRL, ISR, ITA, JAM, JOR, JPN, KEN, KHM, KOR, LAO, LBN, LBR, LKA, LSO, MAR, MDG, MEX, MLI, MMR, MNG, MOZ, MRT, MWI, MYS, NER, NGA, NIC, NLD, NOR, NPL, NZL, PAK, PAN, PER, PHL, POL, PRT, PRY, PSE, ROU, RWA, SEN, SGP, SLE, SLV, SWE, SYR, TGO, THA, TUN, TUR, TWN, TZA, UGA, URY, USA, VEN, VNM, ZAF, ZMB, and ZWE.

[23]Kremer et. al. (2022) exclude countries with populations less than 200,000 which gives a larger sample but makes no qualitative difference to the results.

using the country data weighted by population and we denote the so normalized value of *RGDPO*/*POP* in country $c$ and time $t$ by $y_{c,t}$.

Figure 1 shows the evolution of the quintile boundaries for each year in the dataset with each line showing the indicated fraction of the sample falling below it.

[Figure 1 about here]

The relative constancy of the boundaries between the bottom three quintiles (the 20% and 40% lines) suggests that, relative to the sample mean per capita GDP, there was little change in the lot of those countries in the bottom 40% of the distribution, save but perhaps for the slight improvement implied by the very slow increase in the 40% line beginning in the early-2000s. The 60% line, the gap between the third and fourth quintile, shows less stability than the 20% and 40% lines, and exhibits a potentially important rise that begins in the early-1990s before faltering and then regaining momentum in the early 2000s and then petering out to give way to a decline beginning around 2010. By contrast, the 80% line, the gap between the top quintile and the rest of the sample, is much more volatile than the other three. It rises markedly from the early-1970s through the late-1990s before falling substantially during the 2000s and stabilizing around 2010.

The rises in the 40% and 60% lines and, especially, the fall in the 80% line during the 2000s, together imply a reduction in the spread of the cross-country distribution of per capita income during that decade after three decades of a growing gap between the top and the bottom. That the 60% and the 80% lies move further apart after 2010 suggests a rise in the gap between the top and the middle in the last decade of the sample.

Visual inspection of Figure 1 suggests a lack of catching up behavior over the 1970 to 2019 period. This impression is confirmed by a $\beta$-convergence test. Estimation of a regression of the average growth rate of $y_{c,t}$ from 1970 to 2019 on $\log y_{c,1970}$, that is, of Equation (4), yields an estimated $\beta$ of $-0.104$ with a heteroskedastisity-consistent standard error of 0.181 and so a t-ratio of $-0.574$. This implies p-value of 0.283 for the null hypothesis that $\beta$ is zero against the alternative that it is negative. The consequent failure to reject the null hypothesis means that there is no evidence of $\beta$-convergence over the span of the entire sample period.

There is, however, evidence of the short-term $\beta$-convergence found by Patel et. al. (2021) and Kremer et. al (2022) in the last decade of the data. To show this, we follow Kremer et. al (2022) and estimate the model

$$100(\log(y_{c,t+10}) - \log(y_{c,t}))/10 = \alpha_t + \beta_t \log(y_{c,t}) + \epsilon_{c,t}$$

where $c$ indexes countries, $c = 1, \ldots, 103$, and $t$ indexes years, $t = 1970, \ldots, 2019$, This 10-year-ahead growth regression is estimated separately for each value of the initial year, $t$, from 1970 to 2009 and the estimated $\beta_t$ against $t$ along with the 95% confidence bands, calculated using heteroskedastisity-consistent standard errors, are plotted in Figure 2.

[Figure 2 about here]

The results are very similar to those presented in Kremer et. al. (2022) Figures 2 and A.5 and consistent with those presented in Patel et. al. (2021) Figures 1 and A.1.[24] The point estimate of $\beta_t$ is negative from 1994 onwards and the 95% confidence interval does not contain zero from 1999 onwards implying a finding of the short-term $\beta$-convergence for the initial years 1999 to 2009.[25] In broad terms, this finding mirrors implications of the movements of the quintile boundaries, especially that of the 80% line, in Figure 1 with, for example, the estimates of $\beta_t$ turning statistically-significantly negative as that boundary falls after 2000 and the richest 20% of countries move closer to the world mean, a phenomenon that might be better described as "catching down" rather than as "catching up".

[24] Patel, Sandefur, and Subramanian (2021) take a slightly different approach to the growth regressions, fixing the endpoint and reducing the span over which the average growth rate is calculated as the initial year increases.

[25] This is for a two-tailed test. The more appropriate one-tailed test of the null hypothesis that $\beta_t$ is zero against the alternative that it is negative yields a p-value less that 0.05 from 1999 onwards as well.

The changes in the spread of the distribution seen in Figure 1 are also evident in Figure 3 which plots the sample standard deviation of $y_{c,t}$ in each year.

[Figure 3 about here]

The increase in the spread of the distribution implied by the rise in the 80% line in Figure 1 from the early-1970s through the late-1990s is reflected in the rise of the standard deviation over that period evident in Figure 2. Similarly, the subsequent reduction in the spread during the 2000s suggested by the fall in the 80% line and the rises in the 40% and 60% lines drives a fall in the standard deviation. In the latter part of the graph, the increase in the gap between the 60% and 80% lines is enough to increase the standard deviation despite the rise in the 40% line. As noted above, the decline in the standard-deviation of the cross-country distribution of per capita GDP is indicative of so-called $\sigma$-convergence so the 1970 to 2019 period is characterized by three decades of $\sigma$-divergence before about 15 years of $\sigma$-convergence which then gives way to $\sigma$-divergence towards the end of the period.[26]

Figure 4 considers at the evolution of the cross-country distribution of per capita GDP by presenting kernel density estimates of the distribution for 1970, 1980, 1990, 2000, 2010, and 2019 computed using the Epanechnikov kernel and the Sheather-Jones (1991) plug-in bandwidth.[27]

[Figure 4 about here]

The gray area shows the estimated 95% confidence interval for the population density computed under the assumption of a common population density for 1970, 1980, 1990, 2000, 2010, and 2019. The confidence interval is computed using 1000 replications of a bootstrap that draws (with replacement) samples of 103 observations from the combined 1970, 1980, 1990, 2000, 2010, and 2019 datasets and computes the kernel density estimate for the resampled data, again using the Epanechnikov kernel and the

[26]Qualitatively, the results in Figure 2 are very similar to those presented in Kremer et. al. (2022) in their Figures 2 and A.5.

[27]Silverman (1986) provides a comprehensive treatment of kernel density estimation.

Sheather-Jones (1991) plug-in bandwidth. Then, for each of a finely spaced grid of values in $[0.0, 6.0]$, the 25th largest and the 25th smallest value values of the estimated densities are taken as the upper and lower limits respectively of the confidence interval. Portions of the estimated densities that lie outside the confidence interval thus represent phenomena that are outside of typical behavior for the 1970-2019 period.

For example, the rise in the 80% line in Figure 1 from 1990 to 2000 and its subsequent fall from 2000 to 2010 manifests itself as the estimated density for 2000 being above the upper limit of the confidence interval on $[3.6, 5.0]$. About 4% of the mass of the estimated density in 2000 lies outside of the confidence interval. The rise of the 80% line in Figure 1 from 1970 to 1990 is also evident in Figure 4 as on $[3.1, 3.8]$ the estimated density for 1990 lies above that for 1980 which lies above that for 1970 but, except for a small interval where the 1970 estimated density is outside the confidence interval, that doesn't appear to represent behavior that is atypical for the 1970-2019 period.

Also evident is the improvement in the levels of per capita income in the countries towards the bottom of the distribution suggested by the rises in the 40% and 60% lines in Figure 1, beginning in the early 2000s and the early 1990s as discussed above. In Figure 3 this is seen as reduction in the mass in the estimated densities in the interval $[0.0, 0.7]$ from 1990 to 2000 and beyond. This reduction is, however, however small as it represents only about 3% of the mass of the estimated density in 1990.

The estimated 95% confidence interval in Figure 3 implies multimodality in the cross-country distribution of per capita GDP which, as mentioned above is suggestive of convergence clubs. However, as Pittau et. al. (2010) point out, such multimodality is neither necessary nor sufficient to imply the existence of convergence clubs and hence the absence of convergence in the sense of Equation (1). The lack of sufficiency can be at least partly addressed by considering the extent of the mobility of countries with in the distribution. One way to do that is to note that of the possible 515 crossings from one decade to another (i.e. from 1980 to 1990, for example) of the antimode in the lower bound of the 95% confidence interval (which is almost the same as that in the upper bound), just seven occur between 1970 and 2019.[28] That is, only about 1.4% of the possible movements from the possible basin of attraction associated with the lower mode to that associated with the upper mode or vice versa occur in the sample period.

Another way is examine the intra-distributional mobility is to consider the pairwise Spearman rank correlations of output per capita in the countries between each of the years 1970, 1980, 1990, 2000, 2010, and 2019 as displayed in Figure 5.

[Figure 5 about here]

These correlations are very high: those between adjacent decades exceed 0.95, those between samples two decades apart exceed 0.90, those between samples three decades apart exceed 0.85, and those between samples four decades apart exceed 0.80. As with the examination of the crossings of the antimode in the lower bound of the 95% confidence interval, the impression is that there is very little mobility within the cross-country distribution of per capita income.

In Section 4 we delve more deeply into the movements in the cross-country distribution of per capita GDP by modeling the distribution using the mixture modeling approach used in Pittau et. al. (2010). Consistent with earlier results, we show that the distribution can be modeled as a three-component mixture for each decade from the 1970s to the 2010s with little mobility between the components from one decade to the next, a result that we view as consistent with the presence of convergence clubs.

**3. Empirical evidence from the literature**

In this section we discuss the empirical research that has examined the convergence hypothesis using cross-country data. For reasons given earlier, we focus on tests of the unconditional or absolute versions of the hypothesis although research that has

[28]The seven crossings are Greece, which crosses twice, once upwards and once downwards; and Hong Kong, Ireland, South Korea, Portugal, and Singapore, each of which cross upwards once.

considered the conditional version of the hypothesis is mentioned where it helps the exposition. We begin with a look at $\beta$-convergence tests and extensions that allow for cross-country heterogeneity in order to increase the power of the test against non-convergent alternatives. Next we consider tests based on the $\sigma$-convergence approach, before discussing the literature that considers the entire cross-country distribution and its evolution. The section finishes with a consideration of the time-series approaches to convergence. Looking ahead, the conclusion that we take away from the research examining the absolute convergence hypothesis is that, once the economically important alternative hypothesis of club convergence is explicitly considered, there is strong evidence for a rejection of the view that initial conditions do not matter for long-run outcomes in favor of the view that they do. That is, when broad samples of countries are considered using methods that admit the club convergence alternative hypothesis, the cross-country per capita income data supports the view that there are several convergence clubs rather than the claim that absolute convergence is occurring.[29]

### 3.1 $\beta$-convergence regressions and extensions

The early contributions to the modern convergence literature estimate versions of Equation (4) above and test for $\beta < 0$ i.e. a negative correlation between the log of initial per capita income and its subsequent growth rate. Using samples consisting of broad groups of countries, Baumol (1986), Barro (1991), Dowrick (1992), and others fail to find a negative correlation implying a rejection of the absolute $\beta$-convergence hypothesis. This result invites questions about the obstacles to convergence. Abramovitz (1986) notes that the ability to absorb existing technologies and to attract capital must be present in an economy if catch-up growth is to occur. According to Sachs and Warner (1995) "... open

[29]It is important to bear in mind the identification caveat discussed in Durlauf and Johnson (1995), which applies to all of the empirical growth literature as, with a finite span of data, it is not possible to distinguish between a growth process in which there are multiple steady states, or convergence clubs, and one in which countries transition through different stages of development before reaching a common (stochastic) steady state.

economies tend to converge, but closed economies do not. The lack of convergence in recent decades results from the fact that the poorer countries have been closed to the world." Rodrik (2013) finds $\beta$-convergence in cross-country data on manufacturing output per worker which he attributes to the tradability of manufacturing output as, among other things, that facilitates technology transfer and exposes domestic producers to the discipline of foreign competition. He cites the small manufacturing sectors in low-income countries as a cause of the lack of $\beta$-convergence in per capita incomes.

Other contributions add conditioning variables to control for the effects of possible obstacles to convergence. Barro (1991), Barro and Sala-i-Martin (1992), Dowrick (1992), and Mankiw et. al., (1992), add variables such as rates of capital accumulation, population growth rates, and policy variables, and find a statistically significant negative partial correlation between (the log of) initial per capita income and its subsequent growth rate. Mankiw et. al. (1992) was the first contribution to derive the set of conditioning variables from an explicit growth model and that contribution, along with those of Barro (1991) and Barro and Sala-i-Martin (1992), sparked a voluminous literature testing this so-called conditional $\beta$-convergence hypothesis with wide variety of datasets. The results in these studies are generally affirmed by later work using panel data such as Islam (1995) and Caselli et. al. (1996). Sala-i-Martin (1996) notes that one of the striking results obtained in these studies is the rate at which economies converge to their conditional steady-states is roughly 2 percent a year. Barro (2015) refers to this rate as the "iron law of convergence" and finds further support for it using a broad panel beginning in 1960 and a narrower one beginning in 1870.[30] These results ought not, however, be taken as supportive of the view that catching up is actually occurring. As we have already noted, Dowrick and DeLong (2003, p. 204) describe the presumption that poor countries can catch up to the rich countries simply by adopting the appropriate characteristics of the latter as the “joker in the deck”. They argue that “a moment’s thought will convince

[30] See Abreu et. al. (2005) for a closer look at the "iron law of convergence".

anyone that many of the right-hand-side variables used ... [in conditional convergence tests] ... could never be brought to the mean values found in the industrial core of the world economy in any country that has not already attained the productivity level and socioeconomic structure found in the industrial core."

The restriction of the sample used when estimating Equation (4) to groups of countries that arguably have similar steady states such as the industrialized countries (Abramovitz, 1986; Baumol, 1986; Dowrick and Nguyen, 1989) or the individual states of the United States (Barro and Sala-i-Martin, 1991, 1992) also yields a finding of $\beta$-convergence. Of course, this "local" convergence is consistent with such "countries" lying in the same basin of attraction of the growth process. At least in the case of the industrialized countries, it may also reflect sample selection issues because unsuccessful countries that have not converged were excluded from the group of countries studied, a point made by DeLong (1988) in his critique of Baumol (1986).

More recently, Patel et. al. (2021) and Kremer, et. al (2022) argue that, if growth over relatively short periods of time is considered, there is evidence of unconditional cross-country $\beta$-convergence beginning in the mid- to late-1990s.[31] This short-term $\beta$-convergence is documented in Figure 2 above which is very similar to Figures 2 and A.5 in Kremer et. al. (2022) and consistent with Figures 1 and A.1 in Patel et. al. (2021).[32] As noted in the discussion of Figure 2, this phenomenon might be better described as "catching down" rather than "catching up" as it largely reflects the movement of the richest 20% of countries towards the world mean after 2000 documented in Figure 1. For example, redoing the analysis presented in Figure 2 for the poorest 62 countries in 1970 (i.e. the poorest 60% of our sample of 103 countries) largely removes the short-run $\beta$-convergence phenomenon in that a version of Figure 2 computed using just those 62

[31]See also Roy et. al. (2016).

[32]As Patel et. al. (2001) discuss, the speed of convergence implied by their estimates of $\beta$ is a good deal slower than the "iron law" rate of 2% per year although, as they note, that rate is based on studies of conditional rather than absolute convergence.

countries only rejects the no-convergence null in 1994, 1995, and 1996 using a one-sided 5% test and the point estimate of $\beta$ is positive for most of the post-2000 period.

In general, 10 years is most likely too short of a span of data to reliably assess the presence of convergence or not, regardless of the statistical approach taken. In particular, while convergence as defined in Equation (1) implies that catching up will be observed over some interval, the converse of that proposition is not true as the reduction of the gap in per capita income levels over some interval carries with it no implication of long-run equality of expected per capita income levels. Clearly, the chances of finding an interval over which catching up is evident, even if convergence is not occurring in the sense of Equation (1), will increase as the interval length decreases.

Additionally, there is the known low power of $\beta$-convergence tests against the alternative of club-convergence, an issue first raised by Bernard and Durlauf (1996) who argue that, even if the data generating process has multiple steady states, rejection of the null hypothesis that $\beta = 0$ in favor of the alternative that $\beta < 0$ is not unlikely. Neither Patel et. al. (2021) nor Kremer, et. al (2022) employ the sort of clustering methods that have proven to be useful in the construction of tests with power against the club-convergence alternative and that, when applied to cross-country data, typically reject the convergence hypothesis.[33]

Durlauf and Johnson (1995) seek to improve the power of the $\beta$-convergence test against the club-convergence alternative by using the regression tree method of Brieman et. al (1984) to estimate a version of the human-capital-augmented Solow growth model introduced by Mankiw et. al. (1992) that allows (endogenously determined) subgroups of countries to obey different locally-linear growth equations. They find that the resultant globally non-linear model fits the Mankiw et. al. (1992) data better than the linear model used by those authors, and that the non-linearity in the estimated model is consistent with

[33] A different criticism of the Kremer, et. al (2022) finding of recent short-term $\beta$-convergence is offered by Acemoglu and Molina (2022) who argue that it is due to the lack of the inclusion of country fixed effects in the estimation.

the view that there are multiple basins of attraction in the process describing the evolution of output per capita. Masanjala and Papageorgiou (2004) show that allowing a CES aggregate production function rather than the Cobb-Douglas function employed by Durlauf and Johnson (1995), following Mankiw et. al. (1992), also yields results implying important nonlinearities consistent with multiple regimes in the growth process. Tan (2010) uses a generalization the regression-tree method used in Durlauf and Johnson (1995) and finds similarities in institutional quality and ethnic fractionalization useful in identifying groups of countries with similar long-run behavior. Notably, he finds no such role for geographic factors.

Earlier researchers had also investigated the possibility that different groups of countries might obey different laws of motion and hence exhibit different long-run behavior. Baumol (1986) divides his sample into the industrialized countries, the centrally planned economies and the rest, depending on their status in 1950, in order to explore the possibility a globally nonlinear but locally linear relationship between initial income and subsequent growth with its implication that exists more that one "convergence club". Baumol and Wolff (1988), Grier and Tullock (1989) and Dowrick (1992) all employ exogenous sample splitting methods and find evidence of heterogeneous behavior across different subsamples in their data.

Other researchers have used data-dependent approaches to fit nonlinear versions of the $\beta$-convergence regression. Baumol and Wolff (1988) allow the coefficient on initial income to vary with its square and find that the negative coefficient required by the catch-up hypothesis is evident only for higher values of initial income. Liu and Stengos (1999) report a similar result after estimating a semi-parametric additive partially linear growth $\beta$-convergence regression that allows the coefficients on a measure of human capital accumulation and initial income to vary smoothly with the levels of the respective variables. Chatterji (1992) estimates a version of Equation 4 that includes the square and the cube of log per capita income as explanatory variables and concludes that the implied nonlinear difference equation for per capita income (relative to the US value) has two basins of attraction with countries belonging to one or the other depending on whether their initial income per capita is greater or less than about one-sixth of the US value. Canova and Marcet (1995) employ panel data to identify parameter heterogeneity and find that a dependence of long-run outcomes on initial conditions, contrary to the convergence hypothesis, with differences in initial conditions explaining almost half the cross-country variation in long-run per capita income forecasts, albeit in a small group of countries. Desdoigts (1999) and Kourtellos (2002) use projection pursuit methods to explore the heterogeneity in the growth relationships in the cross-country data. Both find evidence of nonlinearities in the data consistent with the presence of convergence clubs.

Durlauf et. al. (2001) extend the approach of Liu and Stengos (1999) by allowing all of the coefficients of the growth regression implied by the augmented Solow model of Mankiw et. al. (1992) to vary with initial income. This means that, while the model is Solow for any value of initial income, globally it is permitted to exhibit parameter heterogeneity. The resultant estimated model corroborates the findings of Baumol and Wolff (1988) and Liu and Stengos (1999) of a negative coefficient on initial income only for higher values of initial income. Henderson (2010) uses a nonparametric $\beta$-convergence regression to estimate the density of the coefficient on initial income. The estimated density has a mode centered on negative values of $\beta$, implying $\beta$-convergence for some countries (the OECD) and two other modes centered on positive values, implying the absence of $\beta$-convergence for most of the countries in the data. The evidence on heterogeneity is further strengthened by Kourtellos (2011) who allows the coefficients in the $\beta$-regression model to vary with initial literacy rates and life expectancy. He concludes that the initial values of these variables are determinants of long-run outcomes suggesting the existence of multiple steady states in the growth process.[34]

More general nonparametric estimation methods that allow for nonlinearities without needing need to specify the precise variables governing parameter heterogeneity have also been employed. Using a nonparametric local linear estimator, Maasoumi at. al. (2007) estimate growth regressions with the "Solow variables" on the right hand side for both OECD and non-OECD samples and find considerable variation in the relationships between growth and its determinants across the two samples, as well as in deviations from linearity within each sample. The results of Henderson, et. al. (2012) who undertake nonparametric estimation of regression functions in the presence of "irrelevant" regressors, further demonstrate the importance of nonlinearities in growth regressions.[35]

### 3.2 $\sigma$-convergence tests and related approaches

The observation that the dispersion of per capita incomes across economies should fall if catching up is occurring is the motivation for so-called $\sigma$-convergence tests. Durlauf et. al. (2009) discuss some of the econometric issues that arise with this approach. Typically dispersion is measured using the standard deviation or the coefficient of variation of the log of cross country per capita income as in Barro and Sala-i-Martin (1991), Ben-David (1993), Lichtenberg (1994), Carree and Klomp (1997), and Slaughter (1997) but other measures have been employed (Cowell, 1995) and some attention has been given to the effects of spatial relationships on measures of $\sigma$-convergence (Bode and Rey, 2006; Egger and Pfaffermayr, 2009).

More recently, using a group of mostly industrialized countries, Barro (2012) finds a tendency for the cross-country standard deviation of the logs of per capita GDP

---

[34]Related work includes Banerjee and Duflo (2003), Ketteni et. al. (2007), Minier (2007a, 2007b), and Sirimaneetham and Temple (2009) all of which find evidence of nonlinear relationships between growth and its determinants.

[35]Cohen-Cole et. al. (2012) consider the problem of translating the wide variety of evidence on nonlinearities and heterogeneity in the growth process into policy recommendations.

and consumption to decline since the mid-1970s. In a companion result to their findings of short-term $\beta$-convergence, Kremer et. al. (2022) report a falling standard deviation of per capita GDP beginning in the early 2000s as documented in Figure 3 above. As in Figure 3, Kremer et. al. (2022)'s results indicate $\sigma$-divergence from the mid-2010s. Altatas (2023) applies the weak $\sigma$-convergence test of Kong. et.al to PWT data on 72 countries over the 1960 to 2010 period and is unable to reject the null hypothesis of no weak $\sigma$-convergence.

As with standard $\beta$-convergence tests, examination of the evolution of the cross-country standard deviation of the logs of per capita income yields a test that lacks power against the club convergence alternative. There is no intrinsic reason why countries falling in two or more basins of attraction would prevent a decline the dispersion of the cross-country distribution of per capita income. One related test that does have power against that alternative is that formulated by Phillips and Sul (2007a, 2007b, 2009). This test is provided by an algorithm for sorting the data into potential locally convergent groups and so testing the hypothesis of convergence against the club convergence alternative. The algorithm is based on the test of "relative convergence", which is closely related to $\sigma$-convergence tests, exposited in Phillips and Sul (2007a, 2007b, 2009) and outlined above, sorting the data into potential locally convergent groups and testing the hypothesis of convergence against the club convergence alternative.[36]

Phillips and Sul (2009) apply the relative convergence test to data from the US States, the Western OECD countries, and 152 countries from the PWT and, consistent with much of the literature, find evidence of convergence in growth rates, but not in levels, for the US states (1929-1998) and the Western OECD countries (1870-2001 and 1940-2001 but not 1870-1929 nor 1911-1970) nor for the PWT countries over the 1970-

[36]Tomal (2023) provides a useful survey of the theory and applications of the clustering algorithm described in Phillips and Sul (2007, 2009). Kong et. al. (2019) briefly discuss the need for a method of identifying convergence clubs in the context of weak $\sigma$convergence but leave the design and implementation of a suitable algorithm for future research.

2003 period. They apply the sorting algorithm to the PWT data and find evidence of four convergence clubs i.e. they find evidence of local but not global convergence for this dataset. They further show that, while there is no evidence of $\beta$-convergence for the PWT data as a whole, there is such evidence within each of the clubs that they find.[37]

Kindberg-Hanlon and Okou (2020) confirm the existence of convergence clubs using the Phillips and Sul algorithm for a dataset based on the PWT covering the 1970-2018 period. Their comparison of the clubs found when the estimation period is 1970-2000 with those found for the 1970-2018 suggests the possibility that some "emerging market and developing economies" have moved to a club implying a higher long-run level of per capita output. Kindberg-Hanlon and Okou (2020) also examine the cross-club variation in the means of variables that might predict club membership and find that, while education levels, economic complexity, perceived government effectiveness, and initial per capita output are significantly higher in the most developed club, that is not the case for openness and the investment-to-GDP ratio, variables that are widely believed to be positively related to economic development.[38]

### 3.3 Distributional approaches to convergence

A large body of research investigating the convergence hypothesis has focused on the shape and dynamics of the cross-country distribution of per capita income. Much of the motivation for that focus is due to the pioneering work of Quah (1993a, 1993b, 1993c, 1996a, 1996b, 1996c, 1997, 2001) which expounded the "distribution dynamics" approach to studying the role of nonlinearities in economic growth. Originally Markov

[37]Johnson (2020) estimates a local linear version of the relative convergence test of Phillips and Sul (2007a, 2007b, 2009) that allows the key parameter, and hence the forces driving convergence, to vary nonparametrically over time. Apply this approach to data for 18 OECD countries during the 20th century yields evidence of substantial waxing and waning of the forces driving convergence which were weakened by the Great Depression and WW2, and strengthened by the relative stability and increasing economic integration of the post-WW2 period. See also Bergeaud et. al. (2020).

[38]Marrero et. al. (2022) use the Phillips and Sul algorithm to study the convergence of poverty rates in a sample of 104 developing countries over the 1981 to 2015 period and find four convergence clubs for the poverty rate. They fit an ordered logit model to the club memberships that they find and per capita income is a more important determinant of club membership than the inequality of its distribution.

chain methods were used to study the evolution of the cross-country distribution of per capita income but, as Quah recognized, the need to discretize the state space in order to apply those methods changes the probabilistic properties of the data. As a result, more recent applications employ almost exclusively continuous state space methods.[39] These include Andres and Lamo (1995), Fiaschi and Lavezzi (2003a, 2003b and 2007a, 2007b), Johnson (2005), Fotopoulos (2006, 2008), Maasoumi, Racine, and Stengos (2007), Fischer and Stumpner (2008), and Bandyopadhyay (2011, 2014), Barseghyan and DiCecio (2011), in addition to the related research cited in Durlauf et. al. (2005).

Quah's work demonstrates the existence of "twin peaks" in the long-run cross-country income distribution i.e. bimodality. This, and the low degree of within distribution mobility implied by his estimates, is indicative of at least two basins of attraction in the growth process. The bimodality of the cross-country distribution of per capita output, and the low degree of mobility between those modes, found by Henderson et. al. (2008) is consistent with Quah's results. With a similar approach, Anderson (2004) and Anderson et al. (2012) explore changes in the gap between poor and rich countries over time.

More recently attention has turned to the possibility that, at least at times, there are three basins of attraction in the cross country distribution of per capita income. As Pittau et. al. (2010) explain, multimodality is neither necessary nor sufficient for the existence of convergence clubs. Following Paap and van Dijk (1998), Tsionas (2000), Pittau (2005), and Pittau and Zelli (2006a, 2006b), they estimate a finite mixture model of the cross-country income distributions at 5-year intervals for the PWT data for 1960-2000 and find that the distributions can each be well described as the mixture of three components. This result is confirmed by Battisti and Parmeter (2012) who generalize the approach of Pittau et. al. (2010) to exploit the panel nature of their data set. For each year

[39]Bulli (2001) provides a method for rigorously discretizing the state space and finds that doing so strengths the findings on the polarization of cross-country per capita incomes.

1960, 1965, …, 2000, Pittau et. al. (2010) use the estimated parameters of the mixture models to compute the conditional probability that a country belongs to each of the three components and assign countries to the component to which that have the highest probability of belonging. The propensity of countries to change their assigned components over time then provides a measure of within-distribution mobility. Finding little movement between components, Pittau et. al. (2010) interpret the three components that they find as evidence of three basins of attraction in the growth process.

Using the estimated component means, Pittau et. al. (2010) and note that the gap between the implied middle-income and high-income groups widened over the after the early 1970's.[40] Using a related approach, Anderson et. al. (2016) argue that the tendency for many middle income countries to fall back into the poor group is the cause of the widening of the gap between the rich and other countries. This may be an example of the "middle income trap", the label given to the apparent tendency for countries experience abrupt slowdowns following periods of rapid growth (Kharas and Kohli, 2011). Vollmer et. al. (2013b) fit a three-component finite mixture model of the joint distribution of income per capita, educational attainment and life expectancy and argue for the emergence of three "human development clubs" during the 1990's and note the possibility that their result reflects "middle income trap" phenomenon. Importantly, they point out that the apparent trap may represent a transitory stage of development rather than a convergence club, i.e. an example of the identification problem discussed by Durlauf et. al. (2005, pp 622-3).[41] More recently, Patel et. al. (2021) contest the continued existence

[40]Pittau et. al. (2010) also documented a rise in polarization in the cross-country income distribution over the 1960 to 2000 period caused by an increase in the gap between the mean per capita incomes of the countries in the poorest and in the richest of the three components that they find and a decrease in the dispersion of these two groups of countries around their respective component means.

[41]Related work by El-Gamal and Ryu (2012) documents the appearance, disappearance, and reappearance of a "stochastically stable" middle income group over the 1960-2009 period. They note that this could be due to nonstationarity in the observed transition dynamics that they observe, nonlinearities in those dynamics, or a dynamic process of higher than first order. In general, the literature has given little attention to modeling the evolution of the cross country distribution of per capita income by higher-than-first-order processes, an issue taken up by Fiaschi and Johnson (2024).

of the "middle income trap" stating that "[m]iddle-income countries now show the most persistent growth rates, in direct contradiction to the narrative of a middle-income trap".

Several researchers have sought the determinants of the multiple (stochastic) steady states apparently present in the data. The role of "proximate determinants" is studied by Battisti and Parmeter (2013) by using mixture models to model the joint distribution of output per capita, physical and human capital, and TFP. They too find evidence of multiple clusters. Others study the role of possible "deep determinants" of growth in cluster formation. Bloom et. al. (2003) use a mixture model of multiple equilibria to study the role of geographic factors in the determination of per capita incomes large group of countries in 1985. They find an important role for these factors in countries' affinity with the two equilibria that they identify and their economic outcomes. Subsequent work, however, seems to suggest that geographic factors are less important that other factors in convergence dynamics. This difference could reflect the use of only geographic factors by Bloom et. al. (2003). Owen, Videras, and Davis (2009) estimate a finite mixture model for the conditional distribution of growth rates and conclude that the growth process is characterized by multiple regimes with institutional quality being an important determinant of which countries obey which regimes. Canova (2004) orders the OECD countries according to several variables including initial (i.e. in 1950) per capita income, initial human capital, measures of the initial dispersion of income and human capital, geography, and openness and uses a predictive density method to search for evidence of convergence clubs defined by these candidate initial conditions. He finds that the OECD countries can be divided into two clusters defined by the ordering based on initial per capita income with very little mobility between them. After considering a large number and wide variety of possible deep and proximate determinants economic growth as candidates for defining growth clusters, Fiaschi, Lavezzi, and Parenti (2020) conclude that initial conditions define three clusters according to life expectancy in 1960 (a measure of human capital) and the share of Catholics in the population in 1965 (a measure of culture). These results are broadly consistent with Tan's (2010) finding of the importance of similarities in institutional quality and ethnic fractionalization in identifying countries with similar long-run behavior and the lack of such a role for geographic factors, in the context of $\beta$-convergence regressions.

The role of economic openness is highlighted by Epstein et. al. (2003) who apply Quah's distribution dynamics methods to data from 17 OECD countries over the 1870-1992 period. They find considerable persistence within the cross-country distribution of per capita incomes in the pre-1914 and the 1914-1950 periods and, using data for the larger group of 24 countries available for that period, evidence of mobility and some convergence post-1950. A similar difference between the 1870-1914 and post-1950 periods is found by Di Vaio and Enflo (2011) who estimate a mixture model for the growth rate of per capita GDP using data from 1870-2003 for a larger group of countries. Epstein et. al. (2007) return to the role of openness in the convergence or otherwise of per capita incomes by again applying Quah's distribution dynamics methods to data on a broad group of countries over the 1950–1998 period. They find that trade patterns were important in the determination of middle and high income clubs during the "Golden Age" period 1950-1973 as conditioning on trade patterns removes the evidence of convergence clubs found using un conditioned data. By contrast, trade patterns appear to have played little role in the dynamics of the cross-country distribution of per capita income post-1973 despite the unconditional distribution dynamics being very similar in the two periods.

### 3.4 Time series approaches to convergence

In addition to the literature above that focuses on country cross sections or panels, there is a literature studying convergence by exploiting the time series variation in the data. Following Bernard and Durlauf (1995, 1996) much of that literature treats convergence as cointegration although, as Bernard and Durlauf (1996) acknowledge, tests based on the cointegration of cross country income levels are more properly regarded as tests that convergence has occurred, rather than as tests that convergence is occurring. It is possible, for example, that the expected difference in two countries' log per capita incomes at any time includes a deterministic term that goes to zero asymptotically so that convergence as defined in Equation (1) is satisfied despite a contemporary non-zero expected difference in incomes caused by at least one of the countries not having yet reached its steady state. Data generated by such a process may be less likely to yield to rejection of a unit root null hypothesis in the process and so less likely to lead to a conclusion of cointegration if the deterministic term disappears slowly as the difference in incomes will then have a highly persistent component.[42]

As in Bernard and Durlauf (1996), cointegration based tests generally yield results unfavorable to the convergence hypothesis, although transition dynamics could play a role in that conclusion. For example, Pesaran (2007) proposes a test for convergence that involves testing the time-series properties of all possible $N(N-1)/2$ possible pairwise log per capita income deviations for $N$ economies. Using post-war PWT data beginning in 1951 (1540 pairs of countries), 1961 (4851 pairs of countries), and 1971 (5050 pairs of countries) and a variety of ADF tests, Pesaran rejects the hypotheses that the pairs are integrated processes against the non-integrated alternative with frequencies that is often very close to the nominal tests size which, as Pesaran points out, means that the rejections "could have arisen by chance".[43] Testing for linear trends in those pairs for which the unit root null is rejected yields the finding that the number of pairwise deviations that are stationary with a constant mean is never more than 3.8%, a highly unfavorable result as far as the convergence hypothesis is concerned.[44]

---

[42]One attraction of the Phillips and Sul (2007a, 2007b, 2009) relative convergence test discussed above is that it allows for substantial cross-country heterogeneity in the transition dynamics.

[43]Deckers and Hanck (2014) offer a systematic approach to the role of Type I error in this context and confirm Pesaran's (2007) no-convergence conclusion.

[44]Mindful of the low power of the ADF tests, Pesaran (2007) also uses the KPSS test to examine the stationary of the pairwise deviations finding yielding results "marginally more favorable to the convergence hypothesis".

Battisti et. al. (2022) conduct cointegration tests using PWT data on output per worker and labor-augmented TFP ("productivity") for a sample of 103 counties from 1970 to 2017. They conclude that output per worker and productivity are cointegrated for each country so that convergence or otherwise depends on the behavior of productivity. Using the US as the benchmark country, cointegration tests imply substantial divergence in productivity from the US which is taken to represent the "global frontier", driven in large part by the behavior of  countries other than those in the OECD and East Asia. That is, they attribute the observed cross-country divergence per capita incomes to divergence from the global technology frontier.[45] This result is interesting in light of the possibility that rapid technological progress could be the driver of catching-up behavior emphasized by Abramovitz (1986), Baumol (1986), and Marris (1982) among others, who argue that countries behind the technological frontier can move towards it simply through investment in physical capital embodying frontier technologies rather than having to discover anew such technologies for themselves.

To study the role of transition dynamics Oxley and Greasley (1995) add a time trend to the time-series representation for cross-country differences in log per capita income arguing that a negative coefficient indicates the time series version of catching-up is occurring. They find a statistically significant coefficient on this catch up term for the US/UK and US/Australia pairs, using data from the late 19$^{th}$ to the late 20$^{th}$ centuries, but not for the UK/Australia pair which they conclude have converged. Chong et. al. (2008) estimate a non-linear version of the Oxley and Greasley (1995) model for postwar data on the deviations of log per capita income in OECD countries from that in the US. They conclude that transition dynamics alone are not the reason for the rejections of convergence found in time series-tests after not rejecting the unit-root null hypothesis for the 12 cases exhibiting transition dynamics. In contrast, King and Ramlogan-Dobson

[45]See also Comin and Mestieri (2018).

(2011) find that the statistical case against convergence is weakened when breaks in the trend function that captures the transition dynamics are permitted.

Nahar and Inder (2002) also consider the deviations of log per capita income from that in the US modeling it as a polynomial in time and testing the hypothesis that the time-derivative of the polynomial is zero against that alternative that it is positive which implies that the country is catching up to the US. Bentzen's (2005) modifies this approach to allow the rate of convergence to vary over the sample period. Despite highlighting the possible richness of transition paths, this approach specifies the non-convergence null hypothesis as the overly restrictive requirement that a country's deviation of log per capita income from that in the US is an i.i.d. random variable and it is unknown how this test performs when the non-convergence implies a an integrated process as suggested by other researchers. A similar criticism applies to Datta's (2003) time-varying parameter approach to modeling transition dynamics which finds evidence of catching up using post-war data from the OECD countries.

Several researchers have found that allowing structural breaks in the form of mean or trend shifts, either exogenously or endogenously determined, reduces the tendency to not reject the unit root null hypothesis and so makes a conclusion favorably to the convergence hypothesis more likely.[46] These include Carlino and Mills (1993), Oxley and Greasley (1995), Loewy and Papell (1996), Greasley and Oxley (1997), Li and Papell (1999), Strazicich, Lee, and Day (2004), Dawson and Sen (2007), Dawson and Strazicich (2010), King and Ramlogan-Dobson (2011, 2014), Costantini and Sen (2012), and Ghoshray and Khan (2015). In general, the interpretation of these results is not clear even though the breaks are often found to coincide with the Great Depression or the Second World War.[47] If cointegration is found once a structural break is permitted, the

[46]Perron (1989) shows that presence of trend or structural breaks in the data generating process reduces the power of unit-root tests while Stock (1994) surveys the literature on the relationship between structural breaks and unit-root tests.

[47]Using several panel unit root tests and the Bernard and Durlauf (1996) dataset, Fleissig and Strauss (2001) confirm Bernard and Durlauf (1996)'s finding of no convergence among the OECD countries over

conclusion of convergence that follows is then conditional on the occurrence of the break. A model of economic growth that gave a role to such shocks would better inform the empirical work.

Michelacci and Zaffaroni (2000) point out a possible flaw in cointegration-based tests of convergence as log per capita income times series could obey long-memory (or fractionally integrated) processes.[48] While mean reverting, such processes are more persistent than those considered as the alternative hypothesis in the unit-root tests used to test for cointegration and so could lead to erroneous failures to reject a no-cointegration null hypothesis even though convergence is occurring in the sense of Equation (1). While they do not explicitly test for convergence in a long-memory framework, Michelacci and Zaffaroni (2000) do find evidence of long-memory behavior in output data from the OECD countries. Employing several empirical approaches designed to overcome the criticisms of the Michelacci and Zaffaroni (2000) data analysis, Silverberg and Verspagen (2003) reach more agnostic conclusions about the suitability of long-memory representations for per capita output.

Cunado et. al. (2006) test for convergence by examining the (possibly fractional) order of integration of the difference between log per capita income in 14 OECD countries and that in the US. They are unable to reject the hypothesis of a unit root for almost all countries when their data begins in the late 19th century but they find the deviation to be fractionally integrated for almost all of the 14 countries when the analysis is restricted to the post-war period. Silverberg and Verspagen (1999) report similar results. Using post-war data, Dufrénot et. al. (2012) employ long-memory methods to

ver the 1900-1987 period . However, they do find evidence of convergence over the 1948–87 period, a result consistent with the claims of a structure break coincident with WW2. See also the comments in Footnote 37 above.

[48]There is some theoretical support for this possibility with Granger (1980), Michelacci and Zaffaroni (2000), Silverberg and Verspagen (2003) all offering reasons why growth models might imply fractionally integrated processes. By way of contrast however, Lau (1999) shows that integrated and cointegrated processes arise naturally in a wide class of growth models. Granger and Joyeux. (1980) and Baillie (1996) discuss fractional integration and long-memory processes generally.

investigate the behavior of the deviation between log per capita income in a group of 98 developing countries and that in a regional benchmark country (selected as the country in the region with the highest per capita GDP at the end of the sample) and conclude that They find that many of the deviations exhibit non-stationary dynamics, a result that they argue is inconsistent with convergence.

Stengos and Yazgan (2014) follow Pesaran (2007) in avoiding the issue of choosing a benchmark country by considering pairwise convergence in a sample of 139 countries during the post-war period. They conclude against convergence despite allowing the possibility of fractional integration and smooth structural breaks by including a Fourier function of time in the representation for the log per capita income deviations. Stengos et. al. (2018) take a multivariate approach to the estimation of the parameter governing the order of integration and find stronger evidence of mean reversion, and hence convergence, than Stengos and Yazgan (2014).

Another application of Pesaran's (2007) pair-wise time series approach is Beylunioglu et. al. (2020) who use a "maximal clique" clustering approach. Like Hobijn and Frances (2000), who also apply a clustering technique in the based on time-series convergence concepts, they find a large number of very small clusters that they call "convergence clubs". While these results are certainly evidence against global convergence, in both cases, the large number of clusters found makes the convergence clubs interpretation difficult to sustain. It seems possible that the groups found reflect countries between which convergence has occurred rather than those between which convergence is occurring so that a test along these lines that also took into account transition dynamics would produce fewer clusters. One such test is the clustering algorithm of Phillips and Sul (2007a, 2007b, 2009) discussed earlier, which allows for a rich variety of transition behavior and finds four convergence clubs when applied to the post-war PWT dataset (Phillips and Sul, 2009; Kindberg-Hanlon and Okou, 2020).

**3.5 Summary**

This section has discussed the empirical research on the convergence hypothesis using cross-country data, focusing on tests of the unconditional or absolute convergence hypothesis. The literature falls into four broad categories: $\beta$-convergence and its extensions, employing regressions, both linear and nonlinear with initial income as the key explanatory variable; $\sigma$-convergence and other methods analyzing the evolution of the cross-country dispersion of per capita income; methods that consider the entire cross-country dispersion of per capita income and its evolution; and, methods that exploit the time series variation in the data. Our reading of this body of research is that there is overwhelming evidence in favor of the view that convergence in the sense of Equation 1 is not occurring. Rather, it seems that the countries of the world fall into convergence clubs due to the multiple basins of attraction in the growth process as the wide variety of tests that have club-convergence as an alternative hypothesis typically reject the null hypothesis of convergence. Notably, we find recent claims of short-term $\beta$-convergence to be unconvincing objections to the convergence club view. The next section presents a deeper look at the data introduced in Section 2 using the mixture modeling approach of Pittau et. al. (2010).

**4. Another look at the cross-country per capita income distribution**

To underscore the conclusion that the literature strongly supports the view that there are multiple basins of attraction in the economic growth process and hence that multiple convergence clubs are present, we analyze the data introduced in Section 2.2 using finite mixture models to characterize the cross-country distribution of per capita income as in Pittau et. al (2010). The motivation for this approach is that convergence clubs states will manifest themselves as mixture components with little mobility between components so that a test of convergence that has club convergence as an alternative can be constructed using a test of the number of mixture components needed to fit the data and an analysis of the with-in distribution mobility.

Full details of the mixture model approach are given in Pittau et. al (2010). Here we offer a brief summary before turning to the results. The $m$-component mixture model specifies the density of a random variable as $f(x, m, \boldsymbol{\Theta}_m, \boldsymbol{\Pi}_m) = \sum_{j=1}^{m} \pi_j f_j(x, \boldsymbol{\theta}_j)$ where $f_j(x, \boldsymbol{\theta}_j)$ is a probability density function with parameter vector $\boldsymbol{\theta}_j$, for $j = 1, \ldots, m$, $\boldsymbol{\Theta}_m = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_m)$ , the $\pi_j$ are the mixing proportions with $\pi_j > 0$ for $j = 1, \ldots, m$, $\sum_{j=1}^{m} \pi_j = 1$, and $\boldsymbol{\Pi}_m = (\pi_1, \pi_2, \ldots, \pi_m)$. Given $m$, the number of components, and the functional forms of the component densities, $f_j(x, \boldsymbol{\theta}_j)$, the parameters of the model can be estimated by the method of maximum likelihood for which we use the EM algorithm. We use the normal distribution for the component densities so that $f_j(x, \boldsymbol{\theta}_j) = N(x; \mu_j, \sigma_j^2)$ with $\mu_j$ being the mean per capita income in component $j$ and $\sigma_j^2$ the within-component variation in per capita income.

To select the number of components, $m$, we use two information criteria: the AIC with the penalty for an additional parameter increased from the usual two to three, and the BIC.[49] Once $m$ is chosen, the parameter vectors $\boldsymbol{\Theta}_m$ and $\boldsymbol{\Pi}_m$ can be estimated enabling study of the properties of the $m$ component densities.

The $\pi_j$ can be interpreted as the unconditional probability of drawing an observation from component $j$. The conditional probability that $X_i$, observation $i$, is a draw from component $j$ is given by

$$\zeta_{ji} = \frac{\pi_j f_j(X_i, \boldsymbol{\theta}_j)}{\sum_{h=1}^{m} \pi_h f_h(X_i, \boldsymbol{\theta}_h)}. \tag{5}$$

These probabilities can be estimated using the estimated parameters and the estimates used to assign observations to components by assigning observation $i$ to that component with the largest estimated $\zeta_{ji}$. Given a panel of data, mobility can be studied by noting the propensity of the assignment of entity $i$ to change over time.

[49] See Pittau et. al. (2016) for a discussion of the selection of the number of components in mixture models.

We use this approach to estimate individual mixture models for each decade in our dataset. This provides a relatively large number of observations, 1030, for each estimation and allows the estimated number of components and the estimated parameters to vary across decades in a way that the results for any decade do not depend on the data for any other decade, but for the dependence due to that in the data itself. The cost is that the number of components and the parameters are constrained to be constant for an entire decade but, as both evolve slowly or do not change at all, this likely has little downside.[50] For each decade, both the AIC and the BIC select three components for the mixture, consistent with the findings of Pittau et. al. (2010), Battisti and Parmeter (2012), Vollmer et. al. (2013), and Amegual et. al. (2024). We refer to these components in order of their component means as low, middle, and high. The estimated component means, standard deviations, and mixing proportions for each decade for the low, middle, and high components are presented in Figures 6, 7 and 8 respectively.

For each parameter, the estimated 95% confidence interval is also given. This is computed as the parameter estimate ± 1.96 times the appropriate standard error estimated using 1000 bootstrap replications, each of which re-estimates the mixture model after sampling with replacement from the appropriate decade's data to create an artificial sample of 1030 observations. For each parameter estimate, the standard error is computed as the standard deviation of the 1000 values of the parameter estimated with the re-sampled data.

[Figures 6, 7, and 8 about here]

As the first panel in Figure 6 shows, the estimated mean for the low component fell from 0.223 in the 1970s to 0.174 in the 2010s, however the estimated confidence intervals are consistent with a constant mean for the low component across all five decades as the largest confidence interval lower bound (0.164 in the 1980s) is less than

[50]Estimation of a separate model for each year makes little qualitative difference to our conclusions at the cost of far less precision in the parameter estimates.

the smallest confidence interval upper bound (0.225 in the 2010s). This is not the case for either of the other two components. The first panel in Figure 7 indicates that the middle component mean rises from 0.694 in the 1970s to 0.770 in the 1990s before falling to 0.675 in the 2010s. The confidence interval upper bound in the 1970s (0.710) is slightly less than the lower bound in the 1990s (0.713) so, the estimated confidence intervals are not consistent with a constant mean for the middle component across all five decades but, were the lower bound in the 1990s a bit lower, they would be. The first panel in Figure 8 reveals that the estimated mean of the high component exhibits far great fluctuations than the other two component means, rising from 2.66 in the 1970s to 3.46 in the 2000s before falling dramatically in the 2010s to 2.58. The estimated confidence intervals for the mean of the high component in the 1990s and the 2000s do not overlap with those in the other decades, and by a wide margin, while that in the 2010s overlaps with those in the 1970s and 1980s.

The second panels in each of Figures 6, 7, and 8 show the estimated component standard deviations and estimated 95% confidence intervals for the low, middle, and high components respectively. As for the component means, the confidence intervals for the low component standard deviations overlap and so are consistent with a constant standard deviation for the low component across all five decades. The same is true for the estimated standard deviation of the middle component although the intersection of the estimated confidence intervals is a very narrow interval. The estimated standard deviation of the high component exhibits far great fluctuations than those of the other two components, falling from 1.05 in the 1970s to 0.675 in the 1990s before rising to 1.12 in the 2010s . The estimated confidence intervals for the standard deviation of the high component in the 1990s and the 2000s do not overlap with those in the other decades, and by a wide margin, while that in the 2010s overlaps with those in the 1970s and 1980s.

The third panels in each of Figures 6, 7, and 8 show the estimated the mixing proportions and estimated 95% confidence intervals for the low, middle, and high components respectively. For the low and middle components the results are very similar to those for the component means and standard deviations. The confidence intervals for the low component mixing proportions overlap and so are consistent with a constant mixing proportion for this component across all five decades. Similar to the results for the component means, the confidence interval lower bound in the 1970s (0.384) is slightly greater than the upper bound in the 1980s (0.383) so, the estimated confidence intervals are not consistent with a constant mixing proportion for the middle component across all five decades but, were the upper bound in the 1980s a bit higher, they would be. In contrast to the differences in the estimated means and standard deviations for the high component across the five decades, the estimated mixing proportions exhibit much less variation and the estimated 95% confidence intervals overlap are so consistent with a constant mixing proportion for this component across all five decades.

Considering the results by component, they suggest a high degree of stability on the low component with the estimated 95% confidence intervals for the mean, standard deviation, and mixing proportions being consistent with all three parameters being constant across the five decades. The same is almost true also for the middle component although there is some evidence of variation in the means and mixing proportions of that component. By contrast, the results for the high component suggest much greater variation in the parameters across the decades. For that component, the estimated means in the 1990s and the 2000s are much higher than those in the other three decades, which are similar, while the estimated standard deviations in the 1990s and the 2000s are much lower than those in the other three decades, which are similar. That is, the countries in this component moved away from the rest of the distribution and closer to themselves in the 1990s and 2000s before moving back to approximately where they were in the 1980s during the 2010s.  These results are consistent with the fluctuations in the line marking the top quintile in Figure 1 and with the changes in the shape of the cross-country distribution of per capita income in Figure 4 discussed in Section 2.

To assess the within-distribution mobility, we estimate the conditional probability that country $i$ is a draw from each component for each year using Equation (5) and then average the estimated probabilities over the 10 years within each decade. Each country is then assigned to the component to which it has the maximum average estimated probability of belonging. Judged by the magnitude of the conditional probabilities, the general degree of affinity between countries and components is very high with 343 of the total 515 maximum-across-components probabilities being at least 0.90, 428 being at least 0.80, and 465 being at least 0.70. Using this approach, 65 countries are attached to the same component in the 2010s as they are in the 1970s – 62 countries that are attached to the same component in every decade and three countries that briefly change to another before returning to that of the 1970s.[51] There is some upward mobility across components in the from the 1970s to the 2010s as 10 countries rise from the bottom to the middle component over this period[52] and 12 countries rise from the middle to the top component.[53] Downward mobility occurs for 16 countries that fall to a lower component by the 2010s after being in the top or middle component in the 1970s.[54]

Notably, the downward mobility tends to occur earlier in the sample period while the upward mobility is concentrated later in the sample. Of the 22 movements to a higher

[51]Twenty-three countries (Burundi, Benin, Burkina Faso, Bangladesh, Central African Republic, Ethiopia, Haiti, Kenya, Cambodia, Lesotho, Madagascar, Mali, Myanmar, Mozambique, Malawi, Niger, Nepal, Pakistan, Rwanda, Sierra Leone, Togo, Tanzania, and Uganda) remain attached to the low component in all five decades. Sixteen countries (Albania, Brazil, Colombia, Costa Rica, Dominican Republic, Ecuador, Guatemala, Jamaica, Jordan, Lebanon, Morocco, Peru, Philippines, Paraguay, Thailand, and Tunisia) remain attached to the middle component in all five decades. Twenty-three countries (Australia, Austria, Belgium, Canada, Switzerland, Germany, Denmark, Spain, Finland, France, United Kingdom, Greece, China, Hong Kong, Ireland, Israel, Italy, Japan, Netherlands, Norway, New Zealand, Portugal, Sweden, and the United States) remain attached to the high component in all five decades. Bulgaria rises from the middle component to the high component in the 1980s before falling back to the middle component for the remainder of the sample. Cameroon rises from the low component to the middle component in the 1980s before falling back to the low component for the remainder of the sample. Sri Lanka falls from the middle component to the low component in the 1980s before rising back to the middle component for the remainder of the sample.

[52]Bolivia, China, Egypt, Indonesia, India, Laos, Mongolia, Palestine, El Salvador, and Viet Nam.

[53]Argentina, Chile, Hungary, Korea, Malaysia, Panama, Poland, Romania, Singapore, Taiwan, Turkey, and Uruguay.

[54]Côte d'Ivoire, Democratic Republic of the Congo, Ghana, Guinea, Honduras, Liberia, Mexico, Mauritania, Nigeria, Nicaragua, Senegal, Syria, Venezuela, South Africa, Zambia, and Zimbabwe.

component that occur from the 1970s to the 2010s, fully 17 occur in the 2000s or the 2010s.[55] By contrast, of the 16 movements to a lower component that occur from the 1970s to the 2010s, just two occur in the 2000s or the 2010s.[56] At the same time, there is substantial immobility across components in the 2000s and the 2010s as 84 countries do not change components in either decade. The increased upward mobility and decreased downward mobility in the 2000s and the 2010s is possibly a factor behind the short-term $\beta$-convergence reported in Patel et. al. (2021) and Kremer et. al. (2022). However, despite the apparent shift in mobility patterns, even in the 2000s and 2010s, there is clear evidence, in the form of multiple components with little mobility between them, supporting the view that the cross-country distribution of per capita income contains convergence clubs i.e. that convergence as defined in Equation (1) is not occurring notwithstanding the short-term $\beta$-convergence findings.

That said, it is important to temper the conclusion in favor of multiple components and low mobility with the identification caveat discussed in Durlauf and Johnson (1995). As we have just a finite span of data, it is not possible to distinguish unequivocally between a growth process in which there are multiple steady states, or convergence clubs, and one in which countries are in the process of transitioning through different stages of development before reaching a common stochastic steady state and so satisfying the definition of convergence given in Equation (1).

## 5. The future of convergence

For three decades prior to 2008 and the Global Financial Crisis (GFC) the global economy was experiencing sustained, high growth rates, unprecedented in history. Notwithstanding major disparities across countries as documented previously, global growth was solid. Major drivers of this global growth run included openness to trade

[55]Egypt, Indonesia, Korea, Singapore, and Taiwan all move upwards prior to the 2000s.
[56]Both Venezuela and Zimbabwe fall from the middle to the low component, the latter in the 2000s and the former in the 2010s. Venezuela had already fallen from the high component to the middle component in the 1980s.

between virtually all countries, favorable financing conditions marked by historically low interest rates, low and stable commodity prices, and a relatively stable global political environment. It is often stressed that this three-decade period saw China emerging as the main factory of the world achieving growth rates in the double digits and, in doing so, lifting an estimated 770,000 million of its citizens out of poverty.[57]

Since the GFC though, much of the conditions enabling global growth slowly deteriorated as several major shocks hit almost simultaneously leaving behind persistent and deep scars. These shocks include the COVID-19 pandemic – by many accounts the shock of the century striking all countries with most of its punch felt by developing economies not able to apply the kind of accommodative fiscal and monetary policies applied in richer countries to combat the severe effects of shutdowns. Inflationary pressures, due to pandemic-induced supply chain disruptions, along with the rise of commodity and food prices due to the Ukraine war, resulted in the severe tightening of financial conditions and the erosion of investment confidence, especially in developing economies. Perhaps even more concerning are the emerging downside risks to long-run growth and convergence related to climate change and Artificial Intelligence (AI).

As shown in Figure 9, since shortly after the GFC in 2008, growth forecasts have steadily diminished over the medium term. According to the International Monetary Fund (IMF), global growth is projected to slow from 3.5% in 2022 to 2.9% in 2024, which is well below the 2000-2019 average of 3.8%.[58]

[Figure 9 about here]

In addition, recent IMF 5-year projections have growth rates remaining below 3% – the lowest medium-term growth forecast published in any edition of the World Economic Outlook (WEO) since April 1990. According to the October 2023 WEO, "[t]he five

[57]For details, see the joint report by the 2022 World Bank and the Development Research Center of the State Council, P.R. China.

[58]Figure 9, these forecasts, and parts of the discussion, draw from Box 1.1 on p. 26 of *World Economic Outlook, October 2023: Navigating Global Divergences.*

largest emerging markets—Brazil, China, India, Indonesia, and Russia—have contributed about 0.9 percentage point to the decline in medium-term global growth prospects between 2008 and 2023. East Asia and the Pacific's outlook has seen the largest downshift."[59]

The decline in medium-term growth prospects, especially that in developing economies, has concerning implications for the prospects of convergence. The five-year-ahead growth forecasts in the April 2008 WEO implied an expectation of absolute $\beta$-convergence. The estimate of $\beta$ from the April 2008 forecasts is a statistically significant $-0.3021$, suggesting that the poorer countries were expected to grow faster than the rich countries and so close the gap between them. By contrast, the estimate of $\beta$ from the five-year-ahead growth forecasts in the April 2023 WEO is a statistically insignificant $-0.0019$ implying no current expectation of closing the gap between rich and poor countries.[60] These pessimistic prospects come on top of the greater output losses already suffered by poorer countries during the recovery from the pandemic as shown by Brussevich, Liu, and Papageorgiou (2022), a subject which we turn to next.

As the COVID-19 pandemic unfolded, many feared that income inequality between countries would widen on account of unequal access to vaccines, inability to work remotely during lockdowns, and insufficient policy space to enact accommodative policies. Surprisingly, Deaton (2021) and Goldberg (2021) have shown that, contrary to initial priors that the pandemic would increase global income inequality in 2020, there was an acceleration in income convergence. Brussevich, Liu, and Papageorgiou (2022) extend the work of Deaton (2021) by exploring the period of post-pandemic recovery in 2021-2024, using GDP per capita projections from the WEO.

As shown in Figure 10, the acceleration in income convergence in the first year of the pandemic due the slowing in the advanced economies was short lived. Growth in the

[59]*World Economic Outlook, October 2023: Navigating Global Divergences*, p. 26.

[60]The estimated $\beta$ values are taken from Figure 1.1.5 in Box 1.1 on p. 26 of *World Economic Outlook, October 2023: Navigating Global Divergences.*

advanced economies bounced back in subsequent years resulting in a period of divergence in cross-country incomes. This pattern holds across several measures of convergence, including $\beta$- and $\sigma$-convergence. In summary, the pandemic pushed many of the developing economies and low-income countries away from advanced economies as the later countries were able to recover from the shock and revert to pre-pandemic income levels (and higher) fairly quickly while the former countries are recovering more slowly.

[Figure 10 about here]

That said, the challenges faced by the developing economies during the pandemic did not pull them into perpetual crises and many observers and commentators have been surprised by the resilience of some of these countries to the pandemic shock – both on how they handled the health crisis and how their economies managed to start recovering sooner than expected. Unfortunately, the war in Ukraine added to the inflation brought by the pandemic with sharp rises in food and energy prices (Barrett, 2022) and the consequent monetary policy tightening in advanced economies dealt further negative shocks to the developing economies

Perhaps even more threatening to cross-country per capita income convergence are two key emerging challenges: climate change and the AI revolution. Climate change affects negatively the economies of all countries, but the higher frequency, intensity, and duration of natural disasters has more devastating impacts on the economies with smaller GDP – predominantly low-income countries. Moreover, within low-income countries, the poor will likely be the most heavily affected by climate change (Hallegatte and Rozenberg, 2017). Having little influence on the future course of climate, how can these countries cope with the challenges they face as temperatures rise and natural disasters intensify?

Cantelmo, Melina and Papageorgiou (2023) use a DSGE model to study the channels through which weather shocks affect macroeconomic outcomes and welfare in disaster-prone countries. Their main finding is that just accounting solely for more frequent and powerful natural disasters, disaster-prone countries grow on average by 1% less a year than their non-disaster-prone peers. These effects have the potential to trigger a serious divergence process of disaster-prone countries, with the level of their GDP being 115% lower than in non-disaster-prone countries after 30 years. And these are conservative estimates. Insofar climate change continues to increase the magnitude and frequency of natural disasters, such a negative macroeconomic and welfare outcome may become increasingly worse.

Given the devastating impact of natural disasters on low-income countries policy options are very limited: bilateral and multilateral partners would have to significantly scale up their financial contributions following such shocks – even non-concessional loans would not be a viable option given the critical levels of debt distress many low-income countries face.[61] An argument made repeatedly by the authorities of low-income countries is that they expect aid in the form of grants and energy transition funds as they are only the victims of CO2 emissions contributed mostly advanced economies and emerging markets.

A new wave of technological change involving artificial intelligence (AI) – machine learning, robotics, big data, and networks – has led to renewed interest in the impact of pervasive automation on growth, wages, and inequality. While the economic landscape everywhere may shift radically, thus far the literature and the policy debate have focused almost exclusively on advanced economies. Narratives about the impact of a new round of automation on developing economies abound, both optimistic and pessimistic, but there has been very little systematic formal analysis.

Alonso et al. (2022) uses a simple but rich conceptual framework to consider the implications for developing countries of a new wave of technological change that

[61]Early estimates of the climate adaptation cost to developing economies by Margulis and Narain (2010) were in the order of $100 billion. Estimates have since been revised upwards significantly and in the latest report by UN Environmental Programme (2022) reached $340 billion.

substitutes pervasively for labor. It makes simple and plausible assumptions: the AI revolution can be modeled as an increase in productivity of a distinct type of capital that substitutes closely with labor; and the only fundamental difference between the advanced and developing country is the level of TFP. The main result is striking: improvements in the productivity of "robots" drive divergence, as advanced countries differentially benefit from their initially higher robot intensity and developing economies see their predominantly low-skilled jobs being substituted away by robots.

The landscape is likely going to be challenging for developing countries which have hoped for high dividends from a much-anticipated demographic transition.[62] By 2030, more than half of the increase in the global labor force is expected to come from the African continent. This was hailed by policymakers as possibly the continent's big chance to benefit from China's graduating middle-income status (de Carvalho Chamon and Kremer, 2006). The Alonso et al. (2022) findings show that robots may steal these jobs from Africa and unless a drastic shift to productivity gains and education investment is put in place rapidly, Africa's much anticipated demographic transition could yield negative not positive dividends. There is no silver bullet for averting divergence. Developing countries, more urgently than ever before, need to invest in raising aggregate productivity and skill levels so that the labor force be complemented rather than substituted by robots, but of course this is easier said than done. Moreover, our sense is that there is a high the likelihood of the near-term reinforcing of the strength of current club membership, if not the actual emergence of new clubs, given the apparent growing tendencies to geopolitical and geoeconomic fragmentation.

[62]See Korinek and Stiglitz (2019) for a more general argument which emphasizes the need for redistribution to make everyone better off in the face of technical progress, in general.

## 6. Conclusions

This paper has surveyed the empirical literature examining the absolute convergence hypothesis in the context of the cross-country distribution of per capita income. The modern empirical growth literature's initial interest in convergence derives from interest in catching-up behavior as countries with low per capita incomes were hypothesized as being likely to grow more quickly that those with high per capita incomes as the former took advantage of a scarcity- induced higher marginal product of capital and also the possibility of adopting frontier technologies without having to discover them for themselves. More generally, motivated by the standard neoclassical growth model's prediction of a unique, stable steady state (under suitable regularity conditions), the convergence hypothesis states that a country's initial conditions have no effect on (the distribution of) its long-run per capita income level.

The alternative view, that initial conditions matter for countries' long-run per capita income levels is motivated by growth models that, for a variety of reasons, have multiple stable steady states. These models do not exhibit convergence in the sense of the long-run irrelevance of initial conditions as economies obeying such a model will converge to the steady state associated with the basin of attraction in which they begin. This means that long-run per capita incomes will depend on initial conditions as measured by the state variables indicating basin of attraction membership. Economies having initial conditions that are sufficiently similar will lie in the same basin of attraction and so converge locally, forming a "convergence club". The existence of two or more such clubs means that global differences in per capita incomes can persist indefinitely. Taken together, a group of economies obeying such a model will exhibit so called club-convergence with club membership depending on initial conditions.

Researchers have taken several statistical approaches to investigating the veracity of the convergence hypothesis. Early work sought a negative (partial) correlation between countries' initial per capita income levels and their subsequent growth rates or a declining cross-sectional dispersion in the cross-country distribution of per capita income as evidence of a closing gap between rich and poor countries and so indicative of convergence. However, both $\beta$-convergence and $\sigma$-convergence tests, as these two respective approaches have come to be known, have low power against the economically-important club-convergence alternative hypothesis. This issue prompted the use of a variety of clustering approaches to testing the convergence hypothesis. Such approaches interrogate the data seeking groups of countries defined by initial conditions that have similar long-run outcomes. The finding of a several such groups is consistent with the presence of multiply steady states in the growth process and so with convergence clubs while the finding of just one such group, a single convergence club, would be consistent with absolute convergence.

Our reading of the research examining the absolute convergence hypothesis is that, once the alternative hypothesis of club convergence is explicitly considered, there is overwhelming evidence for a rejection of the view that initial conditions do not matter for long-run outcomes in favor of the view that they do and that, as a result, when broad samples of countries are considered, the cross-country per capita income data supports the view that there are several convergence clubs rather than the claim that absolute convergence is occurring.

To underscore this reading we model the cross-country distribution of per capita income in each of the five decades from the 1970s to the 2010s using a finite mixture model. We find that the distribution can be described as a three-component mixture in each decade, consistent with previous research. Using the conditional probabilities of component membership as measures of countries' affinity for each component, we find that most countries are strongly attached to a component and that there is little mobility between components. We interpret these results as supportive of the view that the growth process contains multiple basins of attraction and so as supportive of the club-convergence hypothesis. Notably, the finding of multiple convergence clubs in the 2000s, and 2010s serves as a rebuttal to recent claims of short-term $\beta$-convergence during the 20 years or so prior to the COVID-19 pandemic. That is, we find those claims to be unconvincing objections to the view that long-run per capita income levels are dependent on initial conditions.

In addition to this finding and its implications for the long-run prospects of the countries attached to the low and middle components, looking into the future, we see little reason to be optimistic about the likelihood that poor countries will be able to reduce the gaps between themselves and the rich and middle-income countries going forward in light of the challenges posed by the COVID-19 pandemic, inflation and the associated financial tightening, climate change, and artificial intelligence.

We summarize our conclusions with five lessons from the last 40 years of cross-country convergence empirics:

1. Differences in per capita income between poor and rich countries remain large and persistent.
2. Standard $\beta$- and $\sigma$-convergence tests have low power against the club convergence hypothesis which is implied by many growth models.
3. Tests that admit club convergence as an alternative typically reject the hypothesis of a single convergence club (i.e. convergence) in favor of the alternative of multiple clubs with the modal number of clubs found being three.
4. Despite evidence of short-run $\beta$-convergence in the past 25 years or so, data from 2000s and 2010s reject convergence.
5. In addition to the challenge posed by being in a basin of attraction of the economic growth process with a low mean per capita income, developing countries face formidable headwinds in their quest to catch-up to developed countries.

## References

Abramovitz, M. (1986). “Catching Up, Forging Ahead, and Falling Behind,” *The Journal of Economic History*, 46, 385-406.

Abreu, M., H.L.F. De Groot and R.J.G.M. Florax. (2005). “A Meta-Analysis of $\beta$-Convergence: The Legendary 2%,” *Journal of Economic Surveys*, 19, 389-420.

Acemoglu, D. (2009). *Introduction to Modern Economic Growth*. Princeton University Press.

Acemoglu, D., and C. Molina, (2022). "Comment", *NBER Macroeconomics Annual*, 36, 425-442

Aghion, P. and D.P. Howitt. (2009). Economic Growth. MIT Press.

Alfó, M., Trovato, G., & Waldmann, R. (2008). "Testing for country heterogeneity in growth models using a finite mixture approach," *Journal of Applied Econometrics*, 23, 487–514.

Alonso, C., A. Berg, S. Kothari, C. Papageorgiou, and S. Rehman. (2022). "Will the AI revolution cause a great divergence?" *Journal of Monetary Economics*, 127, 18-37.

Amegual D., G. Fiorentini, and E. Sentana, (2024). *Information matrix tests for multivariate Gaussian mixtures*. mimeo, University of Florence.

Anderson, G. (2004). "Making inferences about the polarization, welfare and poverty of nations: a study of 101 countries 1970-1995," *Journal of Applied Econometrics*, 19, 537-550.

Anderson, G., O. Linton, and T.W. Leo, (2012). "A polarization-cohesion perspective on cross-country convergence," *Journal of Economic Growth*, 17, 49-69.

Anderson, G., M.G. Pittau, and R. Zelli. (2016). "Assessing the Convergence and Mobility of Nations without Artificially Specified Class Boundaries," *Journal of Economic Growth*, 21, 283-304.

Andres, J. and A. Lamo. (1995). *Dynamics of the Income Distribution across OECD Countries*, London School of Economics, Centre for Economic Performance Discussion Paper No. 252.

Alatas, S. (2023). "Revisiting the Solow growth model: new empirical evidence on the convergence debate," *Journal of Economic and Administrative Sciences*, 39, 801-817.

Azariadis, C. (1996). "The Economics of Poverty Traps: Part One: Complete Markets," *Journal of Economic Growth*, 1, 449-496.

Baillie, R.T. (1996). "Long Memory Processes and Fractional Integration in Econometrics," *Journal of Econometrics*, 73, 5-59.

Bandyopadhyay, S. (2011). "Rich States, Poor States: Convergence and Polarisation across Indian States," *Scottish Journal of Political Economy*, 57, 414-436.

Bandyopadhyay, S. (2014). "Twin Peaks – Distribution Dynamics of Economic Growth across Indian States," in *Growth, Inequality and Poverty: Prospects for Pro-Poor Growth*, edited by A. Shorrocks and R. van der Hoeven, Oxford University Press.

Barrett, Phillip. (2022). *How food and energy are driving the global inflation surge*, IMF BLOG. https://www.imf.org/en/Blogs/Articles/2022/09/09/cotw-how-food-and-energy-are-driving-the-global-inflation-surge

Barro, R.J. (1991). "Economic Growth in a Cross-Section of Countries," *Quarterly Journal of Economics*, 106, 407-443.

Barro, R.J. (2015). "Convergence and Modernisation," *Economic Journal*, 125, 911-942.

Barro, R.J. and X. Sala-i-Martin. (1991). "Convergence Across States and Regions," *Brooking Papers on Economic Activity*, 1, 107-182.

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

Barro, R.J. and X. Sala-i-Martin. (2003). *Economic Growth* (Second Edition), McGraw-Hill.

Barseghyan, L. and R. DiCecio. (2011). Cross-country Income Convergence Revisited, *Economics Letters*, 113, 244-247.

Battisti, M., G. Di Vaio and J. Zeira., (2022). "Convergence and Divergence: A new approach, new data, and new results," *Macroeconomic Dynamics*, 26, 945-978.

Battisti, M. and C.F. Parmeter. (2012). "GDP Clustering: A Reappraisal." *Economics Letters*, 117, 837-840.

Battisti, M. and C.F. Parmeter. (2013). Clustering and Polarization in the Distribution of Output: A Multivariate Perspective," *Journal of Macroeconomics*, 35, 144-162.

Baumol, W.J. (1986). "Productivity Growth, Convergence, and Welfare, What the Long-Run Data Show," *American Economic Review*, 76, 1072-1085.

Baumol, W. J., and E. N. Wolff. 1988. “Productivity Growth, Convergence, and Welfare: Reply,” *American Economic Review*, 78, 1155–59.

Ben-David, D. (1993). "Equalizing Exchange: Trade Liberalization and Income Convergence." *Quarterly Journal of Economics*, 108, 653-679.

Bentzen, J. (2005). “Testing for Catching-up Periods in Time-series Convergence. *Economic Letters*, 88, 323-328.

Bergeaud, A., Cette, G., & Lecat, R. (2020). "Convergence of GDP per capita in advanced countries over the twentieth century," *Empirical Economics*, 59, 2509-2526.

Bernard, A. and S.N. Durlauf. (1995). “Convergence of International Output Movements,” *Journal of Applied Econometrics*, 10, 97-108.

Bernard, A. and S.N. Durlauf. (1996). “Interpreting Tests of the Convergence Hypothesis,” *Journal of Econometrics*, 71, 161-173.

Beylunioglu, F. C., Yazgan, M. E., & Stengos, T. (2020). Detecting convergence clubs. *Macroeconomic Dynamics*, 24, 629-669.

Bianchi, M. (1997). “Testing for Convergence: Evidence from Non-Parametric Multimodality Tests,” *Journal of Applied Econometrics*, 12, 393-409.

Bloom, D. E., D. Canning, and J. Sevilla, (2003) “Geography and Poverty Traps,” *Journal of Economic Growth*, 8, 355–78.

Bode, E. and S. J. Rey. (2006). "The Spatial Dimensions and Economic Growth and Convergence," *Regional Science*, 85, 171-176.

Brussevich M., S. Liu and C. Papageorgiou. (2022). *Income convergence or divergence in the aftermath of the COVID-19 shock?* IMF Working Paper 2022/121.

Bulli, S. (2001). “Distribution Dynamics and Cross-Country Convergence: A New Approach,” *Scottish Journal of Political Economy*, 48, 226-243.

Canova, F. (2004). “Testing for Convergence Clubs in Income per Capita: A Predictive Density Approach,” *International Economic Review*, 45, 49-77.

Canova, F., and A. Marcet. (1995). *The poor stay poor: Non-convergence across countries and regions*. C.E.P.R. Discussion Papers, CEPR Discussion Papers: 1265.

Cantelmo, Alessandro, Giovanni Melina and Chris Papageorgiou. (2023). "Macroeconomic outcomes in disaster-prone countries," *Journal of Development Economics*, 161(C).

Carlino, G.A. and L.O. Mills. (1993). “Are U.S. Regional Incomes Converging? A Time-series Analysis.” *Journal of Monetary Economics*, 32, 335-346.

Carree, M. and L. Klomp. (1997). “Testing the Convergence Hypothesis: A Comment,” *Review of Economics and Statistics*, 79, 683-686.

Caselli, F., G. Esquivel, and F. Lefort, (1996). “Reopening the Convergence Debate: A New Look at Cross-country Growth Empirics,” *Journal of Economic Growth*, 1, 363–89

Chatterji, M., (1992). “Convergence Clubs and Endogenous Growth”, *Oxford Review of Economic Policy*, 8, (4) 57-69.

Cho, D. (1996). “An Alternative Interpretation of Conditional Convergence Results,” *Journal of Money, Credit and Banking*, 28, 161-73.

Chong, T.T., M.J. Hinich, V.K. Liew, and K.-P. Lim. (2008). “Time Series Test of Nonlinear Convergence and Transitional Dynamics,” *Economic Letters*, 100, 337-339.

Cohen-Cole, E.B., S.N. Durlauf, and G. Rondina. (2012). “Nonlinearities in Growth: From Evidence to Policy,” *Journal of Macroeconomics*, 34, 42-58.

Comin, D., and Mestieri, M., (2018), "If Technology Has Arrived Everywhere, Why has Income Diverged?", *American Economic Journal: Macroeconomics* 10,137-178.

Costantini, M. and A. Sen. (2012). “New Evidence on the Convergence of International Income from a Group of 29 Countries,” *Applied Economics Letters*, 19, 425-429.

Cowell, F. (1995). *Measuring Inequality*. Harvester Wheatsheaf, London.

Cunado, J., L. Gil-Alana, and D.G. Perez. (2006). “Additional Empirical Evidence on Real Convergence: A Fractionally Integrated Approach,” *Review of World Economics*, 142, 67-91.

Datta, A. (2003). “Time-series Tests of Convergence and Transitional Dynamics,” *Economic Letters*, 81, 233-240.

Dawson, J.W. and A. Sen. (2007). “New Evidence on the Convergence of International Income from A Group of 29 Countries,” *Empirical Economics*, 33, 199-230.

Dawson, J.W. and M. C. Strazicich. (2010). “Time-series Tests of Income Convergence with Two Structural Breaks: Evidence from 29 Countries,” *Applied Economics Letters*, 17, 909-912.

de Carvalho Chamon, Marcos and Michael R. Kremer (2006). "Asian growth and African development," *American Economic Review Papers and Proceedings*, 96, 400-404.

Deaton, Angus. (2021). "Covid-19 and global income inequality," *LSE Public Policy Review*, 1,1-10.

Deckers, T., & Hanck, C. (2014). Multiple testing for output convergence. *Macroeconomic Dynamics*, 18, 199-214.

DeLong, B. (1988). “Productivity Growth, Convergence, and Welfare: Comment,” *American Economic Review*, 78, 1138-1154.

Desdoigts, A. (1999). “Patterns of Economic Development and the Formation of Clubs,” *Journal of Economic Growth*, 4, 305-330.

Di Vaio, G., and K. Enflo. (2011) “Did globalization drive convergence? Identifying cross-country growth regimes in the long run,” *European Economic Review*, 55, 832-844.

Dowrick, S. (1992). “Technological Catch Up and Diverging Incomes: Patterns of Economic Growth 1960-88,” *Economic Journal*, 102, 600-610.

Dowrick, S. and J.B. DeLong. (2003). "Globalization and Convergence", in *Globalization in Historical Perspective*, Bordo M.D., A.M. Taylor and J.G. Williamson, (eds.), University of Chicago Press, Chicago and London.

Dufrénot, G., V. Mignon, and T. Naccache. (2012). “Testing Catching-Up between the Developing Countries: ‘Growth Resistance’ and Sometimes ‘Growth Tragedy’,” *Bulletin of Economic Research*, 64, 470-508.

Durlauf, S.N. and P.A. Johnson. (1995). “Multiple Regimes and Cross-Country Growth Behavior,” *Journal of Applied Econometrics*, 10, 365-284.

Durlauf, S.N., and D. Quah. (1999). “The New Empirics of Economic Growth,” in J.B. Taylor and M. Woodford, (eds.), *Handbook of Macroeconomics*, Vol. 1A, Amsterdam: North Holland Elsevier Science, Chapter 4, 231-304.

Durlauf, S.N., A. Kourtellos, and A. Minkin. (2001). “The Local Solow Growth Model,” *European Economic Review*, 45, 928-940.

Durlauf, S.N., P.A. Johnson, and J.R.W. Temple. (2005). “Growth Econometrics,” in P. Aghion and S.N. Durlauf, (eds.), *Handbook of Economic Growth*, Amsterdam, North-Holland, 555-677.

Durlauf S.N., P.A. Johnson and R.W.J. Temple. (2009). “The Econometrics of Convergence,” in *Palgrave Handbook of Econometrics*, Mills T, Patterson K. (eds.), Applied Econometrics, vol. 2. Palgrave Macmillan: Basingstoke, UK; 1087-1118.

Easterly, W. (2005). "National Economic Policies and Economic Growth: A Reappraisal," in P. Aghion and S.N. Durlauf, eds., Handbook of Economic Growth, Amsterdam, North-Holland, 1, 1015-1059.

Egger, P. and M. Pfaffermayr. (2009). “On Testing Conditional Sigma-Convergence,” *Oxford Bulletin of Economics and Statistics*, 71, 453-473.

El-Gamal, M. and D. Ryu. (2013). “Non-Stationarity and Stochastic Stability of Relative Income Club,” *Review of Income and Wealth*, 59, 756-775.

Epstein, P., P. Howlett and M. Schulze. (2003). “Distribution Dynamics: Stratification, Dolarization, and Convergence among OECD Economies, 1870-1992.” *Explorations in Economic History*, 40, 78-97.

Epstein, P., Howlett, P., & Schulze, M. (2007). "Trade, convergence, and globalisation: The dynamics of the international income distribution, 1950-1998," *Explorations in Economic History*, 44, 100-113.

Feenstra, R. C., R. Inklaar and M. P. Timmer (2015). "The Next Generation of the Penn World Table" *American Economic Review*, 105, 3150-3182, available for download at https://www.rug.nl/ggdc/productivity/pwt/.

Feyrer J.D. (2008). “Convergence by Parts,” *B.E. Journal of Macroeconomics*, 8, 1-35.

Fiaschi, D. and P.A. Johnson. (2024). *Does the cross-country distribution of per capita GDP evolve according to a homogeneous first-order process?*, in progress

Fiaschi, D. and A.M. Lavezzi. (2003a). On the Determinants of Growth Volatility: A Nonparametric Approach,” Discussion Papers 2003/25, Dipartimento di Economia e Management (DEM), University of Pisa, Pisa, Italy.

Fiaschi, D. and A.M. Lavezzi. (2003b). “Distribution Dynamics and Nonlinear Growth,” *Journal of Economic Growth*, 8, 379-401.

Fiaschi, D. and A.M. Lavezzi. (2007a). “Nonlinear Economic Growth: Some Theory and Cross-country Evidence,” *Journal of Development Economics*, 84, 271-290.

Fiaschi, D. and A.M. Lavezzi. (2007b). “Appropriate Technology in A Solovian Nonlinear Growth Model,” *Oxford Review of Economic Policy*, 23, 115-133.

Fiaschi, D., Lavezzi, A. M., & Parenti, A. (2020). "Deep and proximate determinants of the world income distribution," *Review of Income and Wealth*, 66, 677-710.

Fischer, M. and P. Stumpner. (2008). "Income Distribution Dynamics and Cross-region Convergence in Europe," *Journal of Geographical Systems*, 10, 109-139.

Fleissig, A. and J. Strauss. (2001). “Panel Unit-root Tests of OECD Stochastic Convergence,” *Review of International Economics*, 9, 153-162.

Fotopoulos, G. (2006). "Nonparametric analysis of regional income dynamics: The case of Greece," *Economics Letters*, 91(3), 450-457.

Fotopoulos, G. (2008). "European Union Regional Productivity Dynamics: A Distributional Approach," *Journal of Regional Science*, 48, 419-454.

Friedman, M. (1992). “Do Old Fallacies Ever Die?” *Journal of Economic Literature*, 30, 2129-2132.

Furceri, D. (2005). “$\beta$ and $\sigma$-convergence: A Mathematical Relation of Causality,” *Economic Letters*, 89, 212-215.

Galor, O. (1996) “Convergence? Inferences from Theoretical Models,” *Economic Journal*, 106, 1056-1069.

Ghoshray, A. and F.A. Khan (2015). “New Empirical Evidence on Income Convergence,” *Empirical Economics*, 49, 343-361.

Goldberg, P. K., (2021). "The problem with the covid convergence." *Project Syndicate*, Jan 28, 2021. https://www.project-syndicate.org/commentary/covid-convergence-rich-countries-getting-poorer-by-pinelopi-koujianou-goldberg-2021-01.

Granger, C.W.J. (1980). “Long Memory Relationships and the Aggregation of Dynamic Models,” *Journal of Econometrics*, 14, 227-238.

Granger, C.W.J. and R. Joyeux. (1980). “An Introduction to Long Memory Time Series Models and Fractional Differencing,” *Journal of Time Series Analysis*, I, 15-39.

Greasley, D. and L. Oxley. (1995). “A Time-series Perspective on Convergence: Australia, UK and USA since 1870,” *Economic Record*, 71, 259-270.

Greasley, D. and L. Oxley. (1997). “Time-Series Tests of the Convergence Hypothesis: Some Positive Results,” *Economic Letters*, 56, 143-147.

Grier, K. and G. Tullock. (1989). “An Empirical Analysis of Cross-national Economic Growth, 1951-1980,” *Journal of Monetary Economics*, 24, 259-276.

Hansen, B.E. (2000). “Sample Splitting and Threshold Estimation,” *Econometrica*, 68, 575-604.

Hart, P.E. (1995). “Galtonian Regression across Countries and the Convergence of Productivity,” *Oxford Bulletin of Economics and Statistics*, 57, 287-293.

Hallegatte, Stéphane and Julie Rozenberg. (2017). "Climate change through a poverty lens." *Nature Climate Change* 4,250-56.

Henderson, D.J., C.F. Parmeter and R.R. Russell. (2008). Modes, Weighted Modes, and Calibrated Modes: Evidence of Clustering Using Modality Tests," *Journal of Applied Econometrics*, 23, 607-638.

Henderson, D.J., C. Papageorgiou and C.F. Parmeter. (2012). "Growth Empirics without Parameters," *Economic Journal*, 122, 125-154.

Hobijn, B. and P.H. Franses. (2000). "Asymptotically Perfect and Relative Convergence of Productivity," *Journal of Applied Econometrics*, 15, 59-81.

International Monetary Fund, (2003). *World Economic Outlook, October 2023: Navigating Global Divergences*, Washington, DC, available at https://www.imf.org/en/Publications/WEO/Issues/2023/10/10/world-economic-outlook-october-2023

Islam, N. (1995). "Growth Empirics: A Panel Data Approach," *The Quarterly Journal of Economics*, 110, 1127-1170

Islam, N. (2003). "What Have We Learnt from the Convergence Debate?" *Journal of Economic Surveys*, 17, 309-362.

Jerzmanowski, M. (2006). "Empirics of Hills, Plateaus, Mountains and Plains: A Markov-switching Approach to Growth," *Journal of Development Economics*, 81, 357-385.

Johnson, P.A. (2000). "A Non-Parametric Analysis of Income Convergence Across The US States," *Economic Letters*, 69, 219-223.

Johnson, P.A. (2005). "A Continuous State Space Approach to 'Convergence by Parts'," *Economics Letters*, 86, 317-321.

Johnson, P.A. (2020). Parameter variation in the 'log *t*' convergence test. *Applied Economics Letters*, 27, 736-739.

Johnson, S., L. William, C. Papageorgiou and A. Subramanian. (2013). "Is Newer Better? Penn World Table Revisions and their Impact on Growth Estimates," *Journal of Monetary Economics*, 60, 255-274.

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

Jones, L. and R. Manuelli, (1990), "A Convex Model of Equilibrium Growth: Theory and Policy Implications," *Journal of Political Economy*, 98, 1008-1038.

Kelly, M., (1992), “On Endogenous Growth with Productivity Shocks,” Journal of *Monetary Economics*, 30, 47-56.

Kerekes, M. (2012). "Growth miracles and failures in a Markov switching classification model of growth," *Journal of Development Economics*, 98, 167-177.

Ketteni, E., T. P. Mamuneas, and T. Stengos. (2007). “Nonlinearities in Economic Growth: A Semiparametric Approach Applied to Information Technology Data.” *Journal of Macroeconomics*, 29, 555–68.

Kharas, H. and H. Kohli (2011). “What Is the Middle Income Trap, Why Do Countries Fall into It, and How Can It Be Avoided?” *Global Journal of Emerging Market Economies*, 3, 281-289.

Kindberg-Hanlon, G., & Okou, C. I. F. (2020). *Productivity convergence: Is anyone catching up?* The World Bank, Policy Research Working Paper Series: 9378. https://documents1.worldbank.org/curated/en/770631599049513984/pdf/Productivity-Convergence-Is-Anyone-Catching-Up.pdf

King, A. and C. Ramlogan-Dobson. (2011). “Nonlinear Time-Series Convergence: The Role of Structural Breaks,” *Economic Letters*, 110, 238-240.

King, A. and C. Ramlogan-Dobson. (2014). “Are income differences within the OECD diminishing? Evidence from Fourier unit root tests,” *Studies in Nonlinear Dynamics & Econometrics*, 18, 185-199.

Kong J., Phillips P.C.B., Sul D., (2019). "Weak $\sigma$-convergence: Theory and applications?," *Journal of Econometrics*, 209, 185-207.

Kong J., Phillips P.C.B., Sul D.. (2020). "Testing convergence using HAR Inference," *Essays in honor of Cheng Hsiao, Advances in Econometrics*, 41, 25-72, Emerald Group Publishing Limited. DOI: 10.1108/S0731-905320200000041002

Korinek, A. and J. E. Stiglitz. (2021). "Artificial intelligence, globalization, and strategies for economic development," *NBER WP 28453*.

Kormendi, R.C. and P.G. Meguire. (1985). “Macroeconomic Determinants of Growth: Cross-Country Evidence,” *Journal of Monetary Economics*, 16, 141-163.

Kourtellos, A., 2002. “A Projection Pursuit Regression Approach to Cross-Country Growth Data.” *University of Cyprus Discussion Paper*.

Kourtellos, A. (2011). “Modeling Parameter Heterogeneity in Cross-Country Regression Models,” *Discussion Paper 11-2011*, Department of Economics, University of Cyprus.

Kravis, I.B., A. Heston, and R. Summers. (1978). “Real GDP Per Capita for More Than One Hundred Countries,” *Economic Journal*, 88, 215-242.

Krause, M. (2017). “The Millennium Peak in Club Convergence: A New Look at Distributional Changes in The Wealth of Nations,” *Journal of Applied Econometrics*, 32, 621-642.

Kremer, M., J. Willis, and Y. You, (2022). "Converging to Convergence", *NBER Macroeconomics Annual*, 36, 337-412

Lau, S.P. (1999). “I(0) in, Integration and Co-Integration Out: Time Series Properties of Endogenous Growth Models,” *Journal of Econometrics*, 93, 1-24.

Lenkoski, A., T.S. Eicher, and A.E. Raftery. (2014). "Two-Stage Bayesian Model Averaging in Endogenous Variable Models," *Econometric Reviews*, 33, 1-4.

Li, Q. and D. Papell. (1999). “Convergence of International Output: Time series Evidence for 16 Countries,” *International Review of Economics and Finance*, 8, 267-280.

Lichtenberg, F.R. (1994). “Testing the Convergence Hypothesis,” *Review of Economics and Statistics*, 76, 576-679.

Liu, Z. and T. Stengos. (1999). “Non-Linearities in Cross-Country Growth Regressions: A Semiparametric Approach,” *Journal of Applied Econometrics*, 14, 527-538.

Lucas, R. (1988). “On the Mechanics of Economic Development,” *Journal of Monetary Economics*, 22, 3-42.

Maasoumi, E., J.S. Racine, and T. Stengos. (2007). “Growth and Convergence: A Profile of Distribution Dynamics and Mobility,” *Journal of Econometrics*, 136, 483-508.

Mankiw, N.G., D. Romer, and D.N. Weil. (1992). “A Contribution to the Empirics of Economic Growth,” *The Quarterly Journal of Economics*, 107, 407-437.

Marrero, G.A., A. S. Marrero-Llinares, and L. Servén. (2022). *Poverty Convergence Clubs*, ECINEQ Working Paper 2022-619, http://www.ecineq.org/2022/09/06/poverty-convergence-clubs/

Marris, R. (1982). “How Much of the Slow-Down Was Catch-Up?” in *Slower Growth in the Western World*, R.C.O. Matthews (ed.), London: Heinemann

Michelacci, C. and P. Zaffaroni. (2000). “(Fractional) Beta Convergence,” *Journal of Monetary Economics*, 45, 129-153.

Minier, J. (2007a). “Nonlinearities and Robustness in Growth Regressions,” *American Economic Review: Papers and Proceedings*, 97, 388-392.

Minier, J. (2007b). “Institutions and Parameter Heterogeneity,” *Journal of Macroeconomics*, 29, 595-611.

Monfort, M., J. C. Cuestas and J. Ordonez. (2013). “Real Convergence in Europe: A Cluster Analysis,” *Economic Modeling*, 33, 689-694.

Nahar, S. and B. Inder. (2002). “Testing Convergence in Economic Growth for OECD Countries,” *Applied Economics*, 34, 2011-2022.

Nyong, M.O. and O.A. Omobitan. (2013). “Nonlinear Income Convergence and Structural Breaks: Further Empirical Evidence,” *International Journal of Economics and Finance*, 5, 81-88.

Owen, A., J. Videras, and L. Davis. (2009). “Do All Countries Follow the Same Growth Process?” *Journal of Economic Growth*, 14, 265-286.

Oxley, L., and D. Greasley. 1995. “A Time-Series Perspective on Convergence: Australia, UK and USA Since 1870.” *Economic Record* , 71, 259–70.

Paap, R. and H.K. Van Dijk. (1998). “Distribution and Mobility of Wealth of Nations,” *European Economic Review*, 42, 1269-1293.

Margulis, S. and U. Narain. (2010). "The costs to developing countries of adapting to climate change: new methods and estimates," *The Global Report of the Economics of Adaptation to Climate Change Study*. World Bank, Washington DC.

Masanjala, W. and C. Papageorgiou. (2004), “The Solow Model with CES Technology: Nonlinearities and Parameter Heterogeneity,” *Journal of Applied Econometric*s, 19, 171-201.

Papageorgiou, C., and F. Perez-Sebastian. (2004), "Can transition dynamics explain the international output data?," *Macroeconomic Dynamics*, 8, 466-492.

Patel, D., J. Sandefur,and A. Subramanian, (2021), "The new era of unconditional convergence", *Journal of Development Economics*, 152, Article 102687

Perron, P. (1989). “The Great Crash, The Oil Price Shock, And The Unit Root Hypothesis,” *Econometrica*, 57, 1361-1401.

Pesaran, M. H. (2007). "A pair-wise approach to testing for output and growth convergence," *Journal of Econometrics*, 138, 312-355.

Phillips, P.C.B. and D. Sul. (2007a). “Transition Modeling and Econometric Convergence Tests,” *Econometrica*, 75, 1771-1855.

Phillips, P.C.B. and D. Sul. (2007b). “Some Empirics On Economic Growth Under Heterogeneous Technology,” *Journal of Macroeconomics*, 29, 455-469.

Phillips, P.C.B. and D. Sul. (2009). “Economic Transition and Growth,” *Journal of Applied Econometrics*, 24, 1153-1185.

Pittau, M.G. (2005). “Fitting Regional Income Distributions in The European Union,” *Oxford Bulletin of Economics and Statistics*, 67, 135-161.

Pittau M.G. and R. Zelli. (2006a). “Trends in Income Distribution in Italy: A Non-Parametric 90 And A Semi-Parametric Analysis,” *Journal of Income Distribution*, 15, 90-118.

Pittau, M.G. and R. Zelli. (2006b). “Empirical Evidence of Income Dynamics Across EU Regions,” *Journal of Applied Econometrics*, 21, 605-628.

Pittau, M.G., R. Zelli and P.A. Johnson, (2010). Mixture Models, Convergence Clubs, and Polarization,” *Review of Income and Wealth*, 56, 102-122.

Pittau, M.G., R. Zelli and R. Massari, R. (2016). Evidence of Convergence Clubs using Mixture Models. *Econometric Reviews*, 35, 1317-1342.

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

Quah, D. (1993a). “Empirical Cross-Section Dynamics in Economic Growth,” European Economic Review, 37, 426-434.

Quah, D. (1993b). “Galton's Fallacy and Tests of the Convergence Hypothesis,” *Scandinavian Journal of Economics*, 95, 427-443.

Quah, D. (1996a). “Empirics for Economic Growth and Convergence,” *European Economic Review*, 40, 1353-1375.

Quah, D. (1996b). “Twin Peaks: Growth and Convergence in Models of Distribution Dynamics,” *Economic Journal*, 106, 1045-1055.

Quah, D. (1996c). “Regional Convergence Clusters Across Europe,” *CEPR Discussion Paper 1286*

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

Quah, D. (2001). “Searching for Prosperity: A Comment,” *Carnegie-Rochester Conference Series on Public Policy*, 55, 305-319.

Rodrik, D. (2013). “Unconditional Convergence in Manufacturing,” *Quarterly Journal of Economics*, 128, 165-204.

Romer, D., (2018). *Advanced Macroeconomics* (5th Edition) McGraw-Hill, New York

Romer, P.M. (1986). "Increasing returns and long-run growth," *Journal of Political Economy*, 94, 1002–37.

Roy, S., M. Kessler, and A. Subramanian. 2016. “Glimpsing the End of Economic History? Unconditional Convergence and the Missing Middle Income Trap.” Working Paper no. 438, *Center for Global Development*, Washington, DC.

Sachs, J. D. and A. M. Warner. (1995), “Economic Reform and the Process of Global Integration.” *Brookings Papers on Economic Activity* 1, 1–118.

Sala-i-Martin, X. (1996). “The Classical Approach to Convergence Analysis,” *Economic Journal*, 106, 1019-1036.

Sheather, S. J. and M. C. Jones, “A Reliable Data-based Bandwidth Selection Method for Kernel Density Estimation,” *Journal of Royal Statistical Society* B, 53, 683–90, 1991.

Silverberg, G. and B. Verspagen. (1999). “Long Memory in Time Series of Economic Growth and Convergence,” *Maastricht University*

Silverberg, G. and B. Verspagen. (2003). “Long Memory and Economic Growth in the World Economy Since the 19th Century,” *Processes with Long Range Correlations: Theory and Applications*, 270-285.

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

Sirimaneetham, V. and J.R.W. Temple. (2009). “Macroeconomic Stability and the Distribution of Growth Rates,” *World Bank Economic Review*, 23, 443-479.

Slaughter, M.J. (1997). "International Trade and Per Capita Income Convergence: A Difference-In-Difference Analysis," *Journal of International Economics*, 55, 203-228.

Stengos, T. and M.E. Yazgan. (2014). “Persistence in Convergence,” *Macroeconomic Dynamics*, 18, 753-782.

Stengos, T., M, E. Yazgan., and H. Özkan, (2018). "Persistence in Convergence and Club Formation," *Bulletin of Economic Research*, 70, 119-138.

Stock, J. H. (1994). "Unit roots, structural breaks and trends," *Handbooks in Economics, vol. 2*. Engle R. F., McFadden D. L.,(eds.) , Amsterdam; London and New York: Elsevier, North-Holland.

Strazicich, M., J. Lee, and E. Day. (2004). “Are incomes converging among OECD countries? Time series evidence with two structural breaks,” *Journal of Macroeconomics*, 26, 131-145.

Sul. D. (2019). *Panel Data Econometrics: Common Factor Analysis for Empirical Researchers*, Routledge, London

Summers, R. and A. Heston. (1988). “A New Set of International Comparisons of Real Product and Price Levels Estimates for 130 Countries: 1950-1985,” *Review of Income and Wealth*, 26, 19-66.

Summers, R., and A. Heston. (1991). "The Penn world table (mark 5): An expanded set of international comparisons, 1950-1988," *Quarterly Journal of Economics*, 106, 327-368

Tan, C.M. (2010). “No One True Path to Development: Uncovering The Interplay Between Geography, Institutions and Ethnic Fractionalization in Economic Development,” *Journal of Applied Econometrics*, 25, 1100-1127.

Temple, J.R.W. (1999). “The New Growth Evidence,” *Journal of Economic Literature*, 37, 112-156.

Tomal, M. (2023). A review of Phillips-Sul approach-based club convergence tests. *Journal of Economic Surveys*, 00, 1–32. https://doi.org/10.1111/joes.12563

Tsionas, E.G. (2000) “Regional Growth and Convergence: Evidence from the United States,” Regional Studies, 34, 231-238.

United Nations Environmental Programme (UNEP). (2022). "Too little, too slow," *Adaptation Gap Report*. UN Press, New York NY.

Vollmer, S., H. Holzmann, and F. Schwaiger. (2013a). “Peaks vs Components,” *Review of Development Economics*, 17, 352–364.

Vollmer, S., H. Holzmann, F. Ketterer, S. Klasen, and D. Canning. (2013b). “The Emergence of Three Human Development Clubs,” *PLoS ONE* 8: e57624. https://doi.org/10.1371/journal.pone.0057624.

Weil, D. (2014). *Economic Growth* (3rd Edition), Routledge, London and New York

World Bank and the Development Research Center of the State Council, P.R. China. (2022). *Four Decades of Poverty Reduction in China*. World Bank, Washington DC.

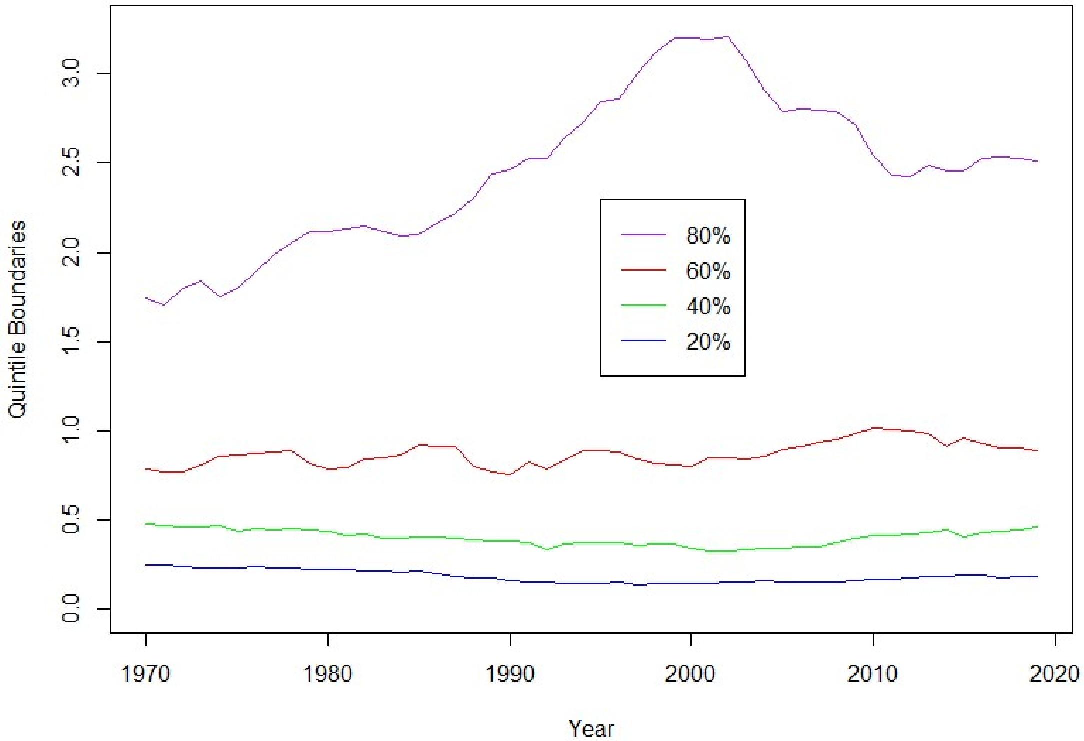
World Economic Outlook (October, 2019). *Global Manufacturing Downturn, Rising Trade Barriers*. International Monetary Fund, Washington DC.

World Economic Outlook (October, 2021). *Recovery During A Pandemic*. International Monetary Fund, Washington DC.

World Economic Outlook (October, 2023). *Navigating Global Divergences*. International Monetary Fund, Washington DC.

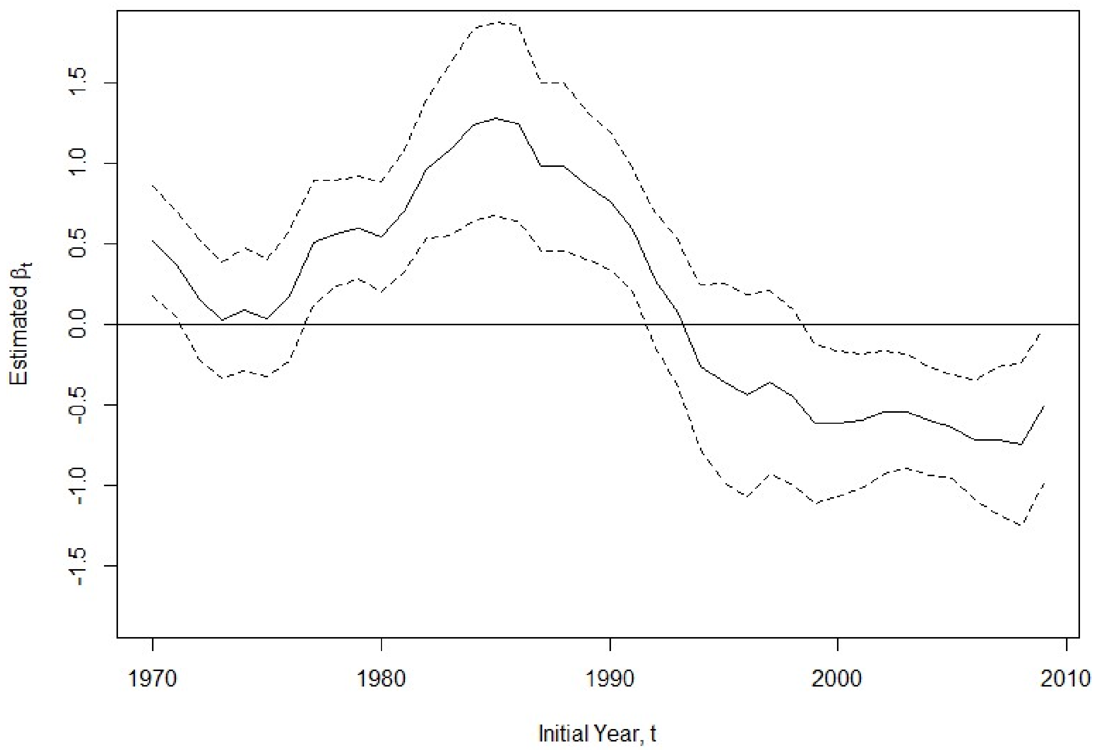
Young, A., M. Higgins, and D. Levy. (2008). "Sigma Convergence Versus Beta Convergence: Evidence from U.S. County-Level Data," *Journal of Money, Credit and Banking*, 40, 1083-1093.

**Figure 1: Quintile Boundaries**



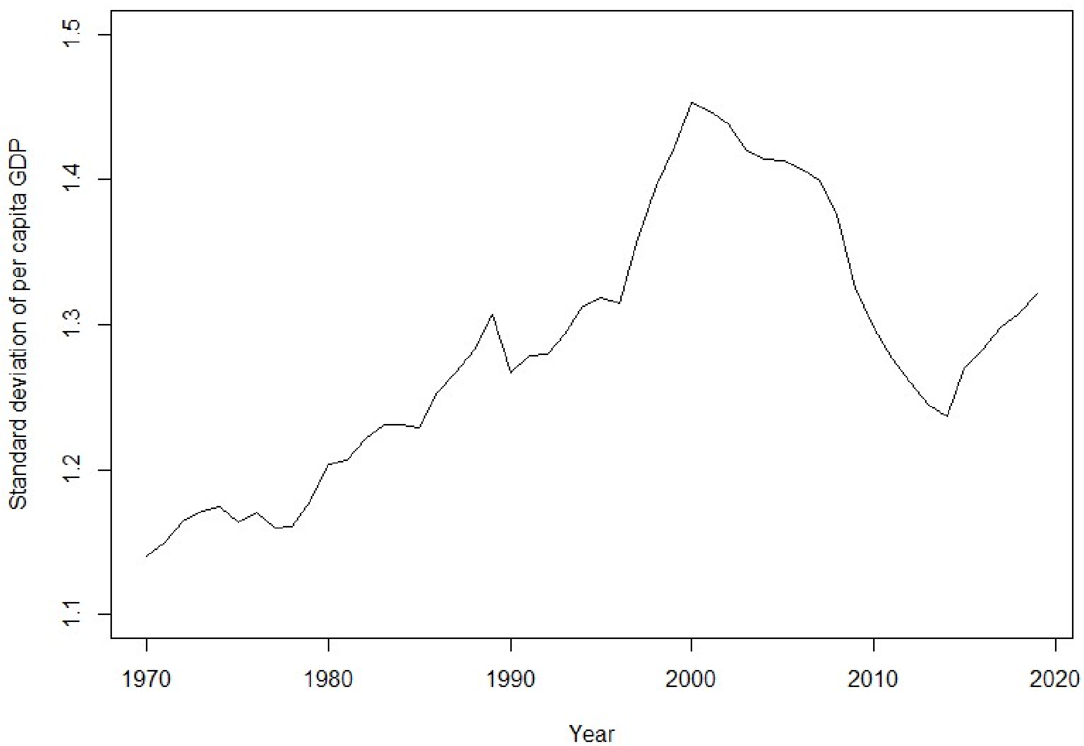

This graph shows the boundaries of the quintiles (the fraction of the sample falling below each line is indicated) of the distribution of $y_{c,t}$, per capita income in country $c$ at year $t$, normalized by the year $t$ sample average of per capita income, for each year in the sample. Authors' calculations using data from Feenstra, Inklaar, and Timmer (2015).

Figure 2: Estimated values of $\beta_t$ and 95% confidence bands



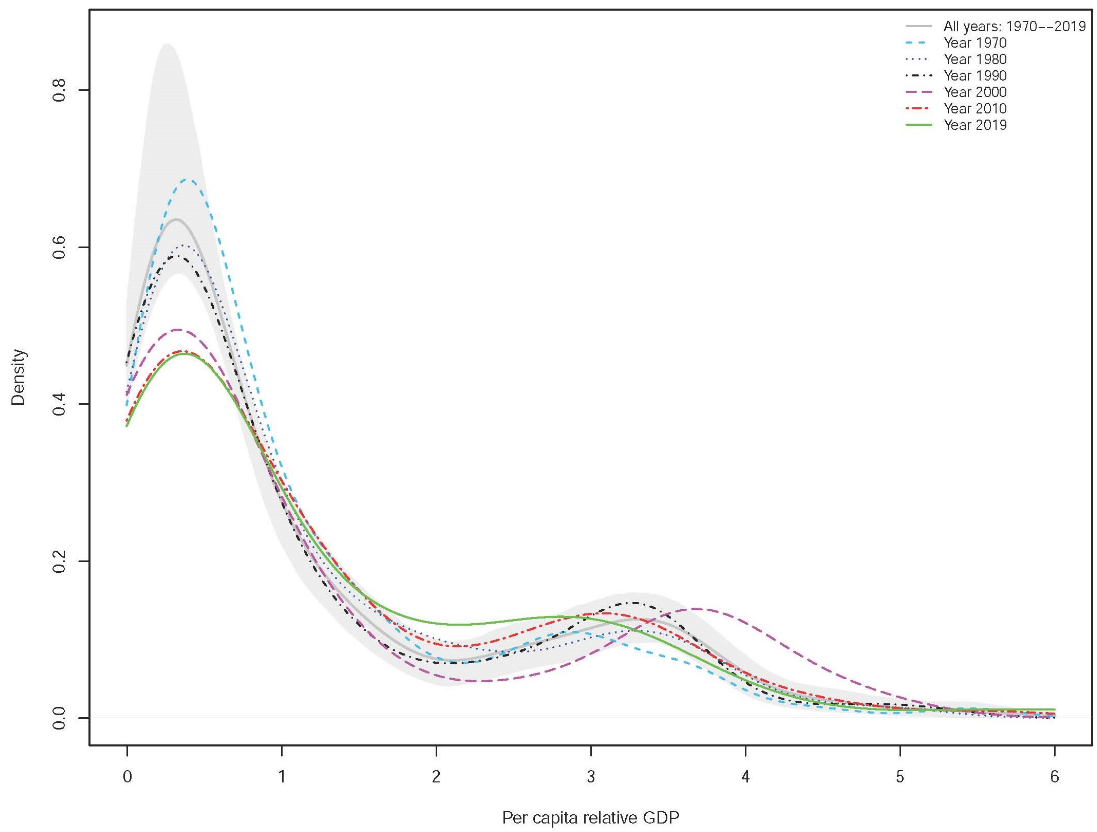

This chart plots the estimated values of $\beta_t$ in the model $100(\log(y_{c,t+10}) - \log(y_{c,t}))/10 = \alpha_t + \beta_t \log(y_{c,t}) + \epsilon_{c,t}$ where $y_{ct}$ is per capita income in country $c$ at year $t$, normalized by the year $t$ sample average of per capita income, from 1970 to 2009. The 95% confidence bands are calculated using heteroskedastisity-consistent standard errors. Source: Feenstra, Inklaar, and Timmer (2015) and authors' calculations.

Figure 3: σ-divergence and σ-convergence 1970 to 2019



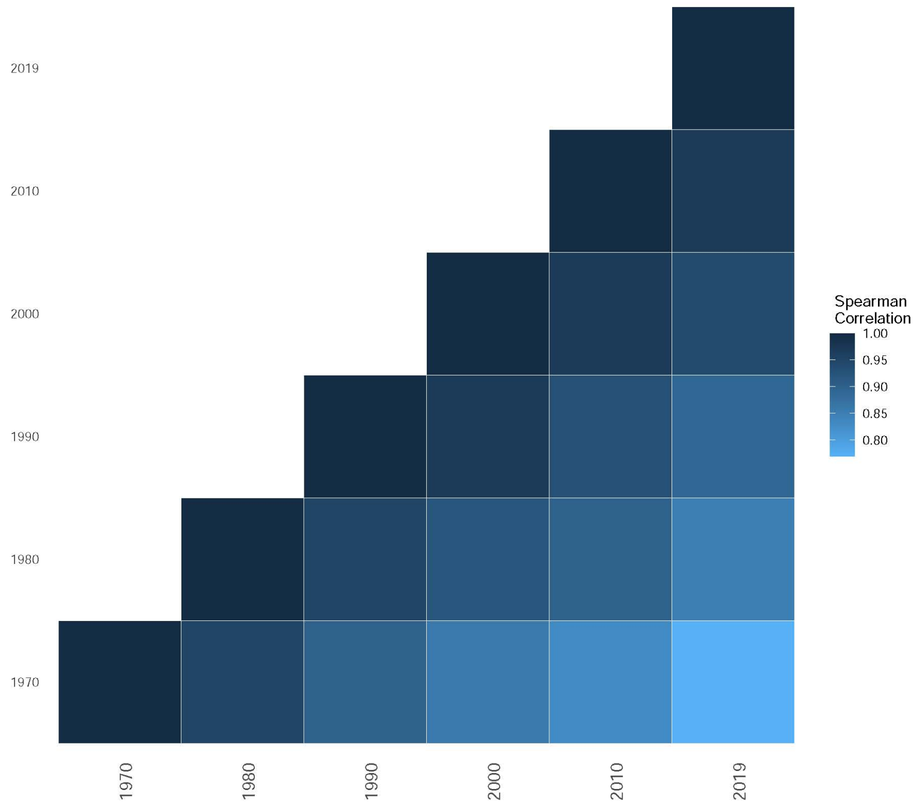

This chart shows the evolution of the sample standard deviation of $y_{c,t}$, per capita income in country $c$ at year $t$, normalized by the year $t$ sample average of per capita income, from 1970 to 2019. Source: Feenstra, Inklaar, and Timmer (2015) and authors' calculations.

Figure 4: Estimated Cross-Country Distributions of Per Capita Income



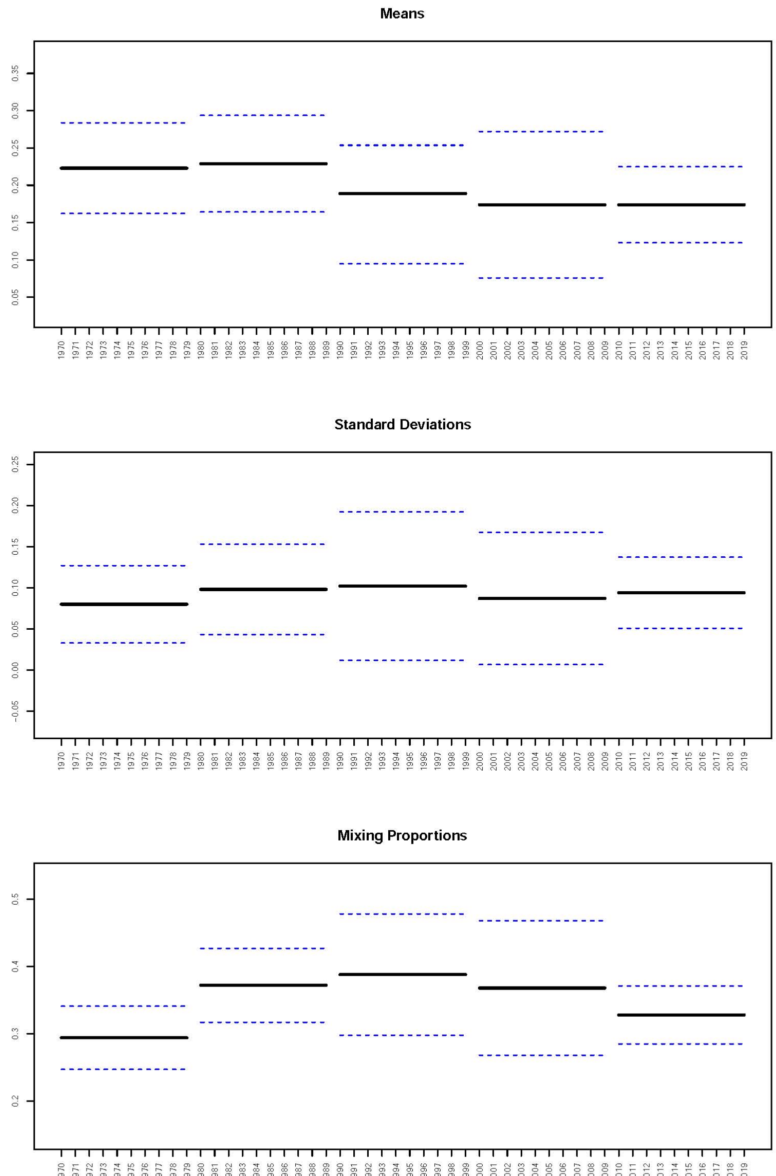

This chart shows kernel density estimates of the distribution of $y_{c,t}$, per capita income in country $c$ at year $t$, normalized by the year $t$ sample average of per capita income, for 1970, 1980, 1990, 2000, 2010, and 2019, and a 95% confidence interval. The estimates are computed using the Epanechnikov kernel and the Sheather-Jones (1991) plug-in bandwidth. The confidence interval is computed using a bootstrap procedure described in the text. Source: Feenstra, Inklaar, and Timmer (2015) and authors' calculations.

Figure 5: Cross-Decade Rank Correlations



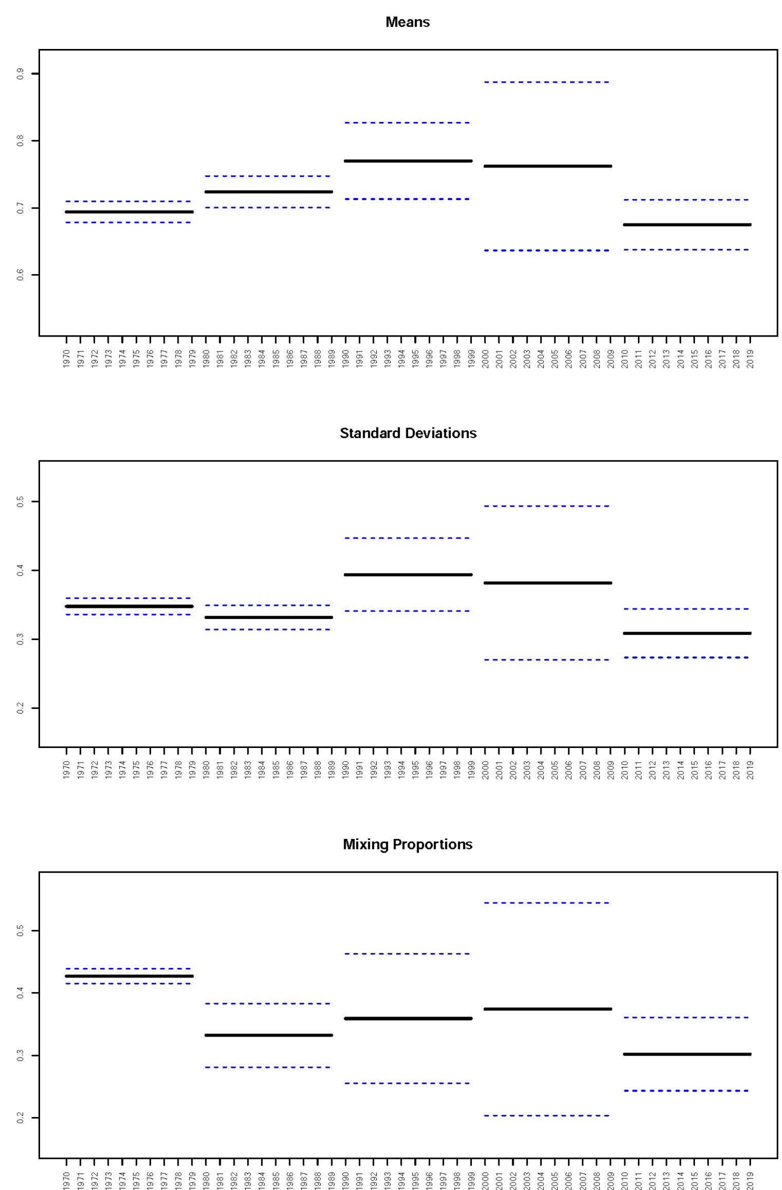

This chart shows the pairwise Spearman rank correlations of $y_{c,t}$, per capita income in country $c$ at year $t$, normalized by the year $t$ sample average of per capita income, between 1970, 1980, 1990, 2000, 2010, and 2019. Source: Feenstra, Inklaar, and Timmer (2015) and authors' calculations.

Figure 6: Low Component Mixture Model Estimates



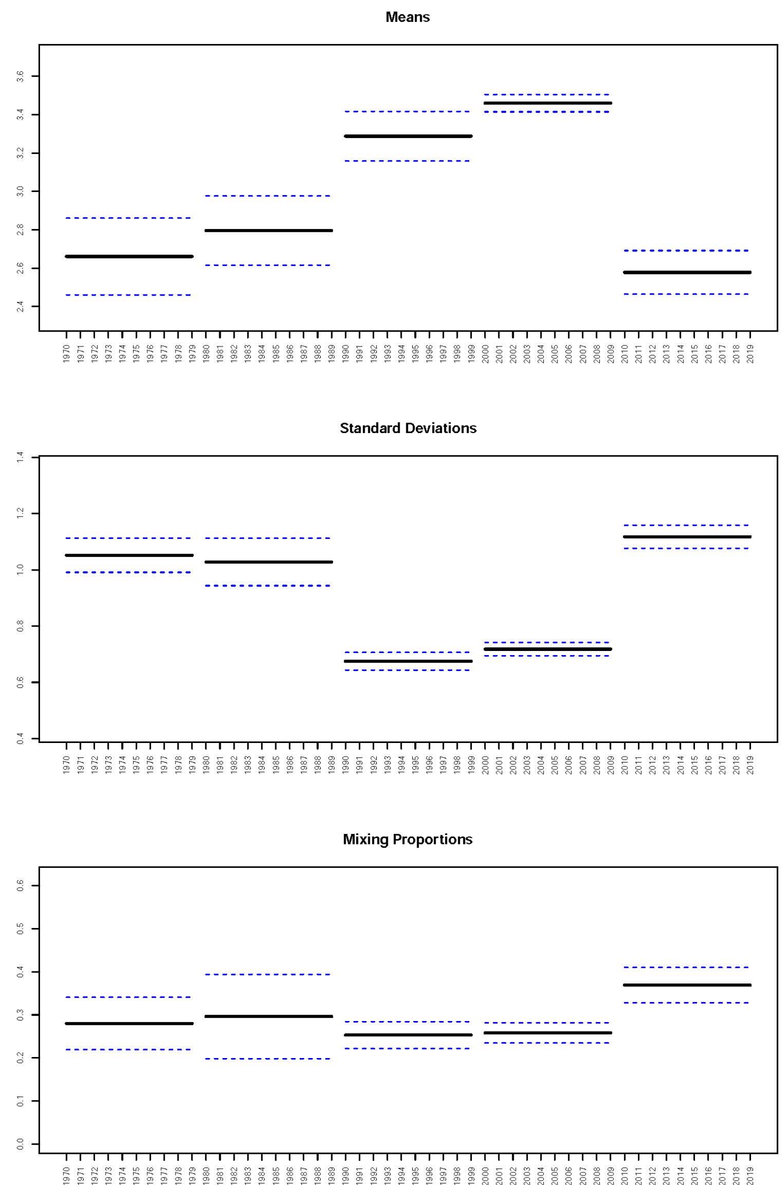

This chart shows the estimated parameters and 95% confidence intervals for the low component of the 3-component mixture model estimated for the distribution of $y_{c,t}$, per capita income in country $c$ at year $t$, normalized by the year $t$ sample average of per capita income. The estimated model specifies the density of $y$ as $\sum_{j=1}^{3} \pi_j N(y; \mu_j, \sigma_j^2)$ where the $\pi_j$ are the mixing proportions with $\pi_j > 0$ for $j = 1, 2, 3$, $\sum_{j=1}^{3} \pi_j = 1$ and $N(x; \mu, \sigma^2)$ denotes the normal distribution with mean $\mu$ and variance $\sigma^2$. The model is estimated separately for each decade from the 1970s to the 2010s by the method of maximum likelihood using the EM algorithm. The confidence intervals are estimated using the bootstrap procedure described in the text. Source: Feenstra, Inklaar, and Timmer (2015) and authors' calculations.

Figure 7: Middle Component Mixture Model Estimates



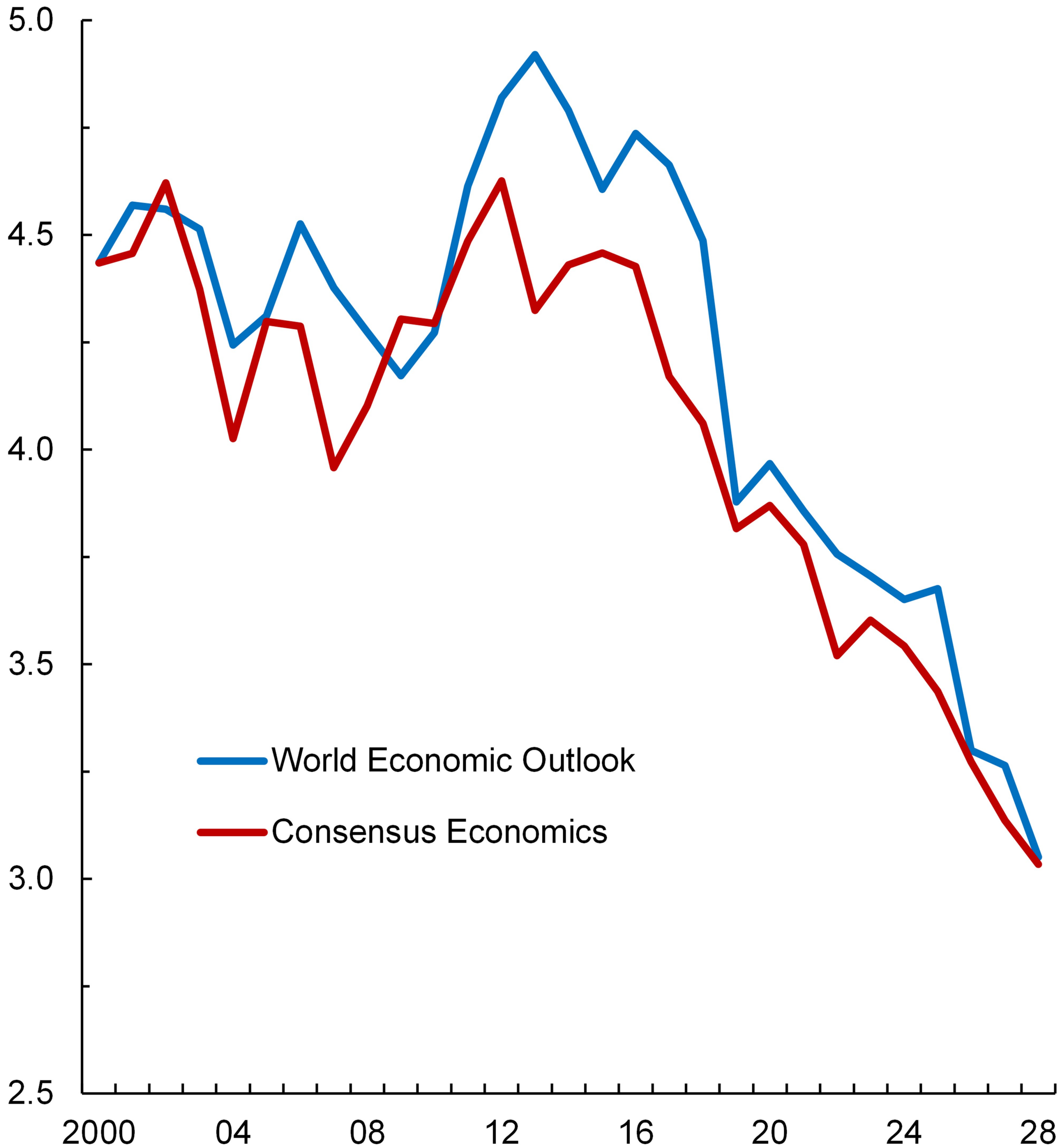

This chart shows the estimated parameters and 95% confidence intervals for the middle component of the 3-component mixture model estimated for the distribution of $y_{c,t}$, per capita income in country $c$ at year $t$, normalized by the year $t$ sample average of per capita income. The estimated model specifies the density of $y$ as $\sum_{j=1}^{3} \pi_j N(y; \mu_j, \sigma_j^2)$ where the $\pi_j$ are the mixing proportions with $\pi_j > 0$ for $j = 1, 2, 3$, $\sum_{j=1}^{3} \pi_j = 1$ and $N(x; \mu, \sigma^2)$ denotes the normal distribution with mean $\mu$ and variance $\sigma^2$. The model is estimated separately for each decade from the 1970s to the 2010s by the method of maximum likelihood using the EM algorithm. The confidence intervals are estimated using the bootstrap procedure described in the text. Source: Feenstra, Inklaar, and Timmer (2015, and authors' calculations.

Figure 8: High Component Mixture Model Estimates



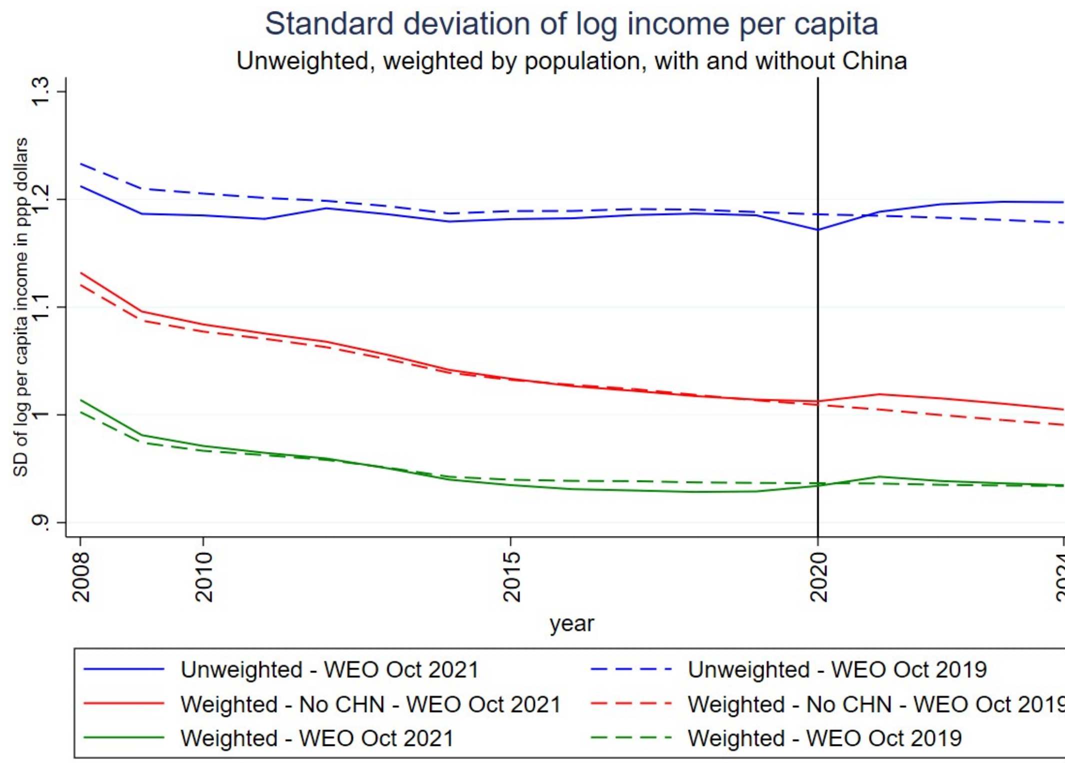

This chart shows the estimated parameters and 95% confidence intervals for the high component of the 3-component mixture model estimated for the distribution of $y_{c,t}$, per capita income in country $c$ at year $t$, normalized by the year $t$ sample average of per capita income. The estimated model specifies the density of $y$ as $\sum_{j=1}^{3}\pi_j N(y;\mu_j,\sigma_j^2)$ where the $\pi_j$ are the mixing proportions with $\pi_j > 0$ for $j = 1,2,3$, $\sum_{j=1}^{3}\pi_j = 1$ and $N(x;\mu,\sigma^2)$ denotes the normal distribution with mean $\mu$ and variance $\sigma^2$. The model is estimated separately for each decade from the 1970s to the 2010s by the method of maximum likelihood using the EM algorithm. The confidence intervals are estimated using the bootstrap procedure described in the text. Source: Feenstra, Inklaar, and Timmer (2015) and authors' calculations.

Figure 9: Five-Year-Ahead Growth Projections




Source: World Economic Outlook (WEO), October 2023, Figure 1.1.1. The predicted variable is real GDP growth. The years on the horizontal axis refer to the year for which the forecast is made, using the April WEO five years prior, such that, for example, the 2028 forecast is based on the April 2023 WEO, and so on. The red line depicts the mean of the Concensus Economics forecasts.

Figure 10: Standard deviation of log income per capita




Source: World Economic Outlook (WEO) October 2019 and October 2021. For Oct. 2021 WEO, 2021 is an estimate and from 2022 onwards are projections. For Oct. 2019 WEO, 2019 is an estimate and from 2020 onwards are projections.