The Next Hundred Years of Growth: Growth and Convergence

Richard Startz*

Revised May 2017

Abstract

I use a Bayesian Markov-switching model to forecast world GDP per capita over the next 100 years. The switching model estimates the probability, for each country in the world, that a country is currently on a path to converge to the world frontier, as well as the probability that a country that is not currently converging will switch to a convergent path. Forecasts depend on both the rate of growth in income in countries at the world frontier and the rate at which other countries converge to that frontier. Convergence is a major source of growth. The contribution of convergence to growth is largely due to countries that have already begun to converge, rather than countries switching from non-convergent to convergent paths in the future. I forecast world income per capita to grow over the next hundred years at an annual rate of 2.7 percent. World income will be almost four times today's U.S. income, and inter-country inequality will be reduced. Almost everyone will have income above today's U.S. income, but only very small fraction will have income at or above future U.S. income.

JEL codes: D04, C11

Keywords: growth, convergence, Markov-switching, Bayesian estimation

---

* Department of Economics, University of California, Santa Barbara, startz@ucsb.edu. I am grateful to Jeremy Piger for code and comments and to Shelly Lundberg and the UCSB Econometrics Working Group for comments.

How well-off will the world be in 100 years, as measured by per capita GDP? The answer depends on growth in the world productivity frontier and on the rate at which poorer countries move toward that frontier. The latter is key: Which countries are already on a convergent path; What is the probability that non-convergent countries switch to the convergent path; and What is the rate of convergence once on that path?

Three basic points are made. The first, the answer to the substantive questions, is that the world is likely to have much, much higher income 100 years from now than is true today and that in large part this will be due to the convergence of poorer countries to the frontier. The second point is that the high future growth rate comes mostly from forecasting that countries that have already started along a convergent path will continue to converge—slowly. The third point is that the emphasis placed here on convergence follows from looking at recent data.

The basic idea is simple: At a point in time a country may be on a path that will eventually lead it to converge to the world frontier, or it may be on a path in which the gap from the frontier is not closing, or possibly on a path in which the country is falling further below the frontier. I model each path as a state of a Markov process, estimating the probability that each country is currently on a given path and the probabilities that countries switch from one path to another. Add to this estimates of how fast a gap closes when on the convergent path and an estimate of frontier growth, and projecting probabilistic future paths for world income is straightforward.

To motivate the empirical work, I offer a variant on the theoretical model in Lucas' (2000) "Some Macroeconomics for the 21st Century." Lucas answers these questions with a "reduced-form," calibrated model. In contrast, a large econometric literature has shown that, looking backwards, some countries appear to be converging toward the world frontier but others appear to be non-convergent, and has attempted to sort out causality for determining which countries are converging.[1] I take an approach somewhere in the middle. I maintain Lucas' reduced-form set-up and also maintain his assumption that *eventually* countries begin to converge to the frontier. Since I estimate the probability of a being on a convergent path, the data decides whether convergence is quantitatively important or not.

My approach picks parameters in a manner that is mid-way between Lucas' calibration and the unrestricted estimation used in much of the literature. I employ a Bayesian framework, positioned part way between calibration and purely data-driven estimates. The Bayesian framework allows one to impose assumptions that limit the range of parameters. In particular, this helps the model pick out low frequency rather than business cycle changes. The Bayesian framework forces one to be explicit about what assumptions are made and I specify priors that are informed by what we know about growth. At the same time, the priors are loose enough that the data chooses parameter estimates within a broad range.

The clear projections made here may seem somewhat surprising given that the empirical literature on convergence has been less than conclusive. Or, as Durlauf (2005) writes, "…the study of growth…cannot easily be distilled into a consensus." The empirical method deployed here is different from what has been used in the past, but what I suspect is more important is that the world has changed since the growth convergence literature was most active. The last 25 years show more evidence of countries converging toward frontier income than did the

[1] See Temple (1999) for a detailed discussion and many references. See also, Jones and Vollrath (2014, section 3.2) and Acemoglu (2009, sections 1.3-1.5) and Barro and Sala-i-Martin (1995, Chapter 12).

earlier postwar period. The estimates here reflect this change, and the increased tendency toward convergence can be seen in the raw data as well. Projections forward are based on postwar history. In particular, the estimate of the probability that a country which has begun to converge falls off that path—which is quite small—largely reflects the last quarter century. So a reader who believes that recent history is an aberration will find the forecasts here too optimistic.

To see why focusing on convergence matters, I offer two not-very-sophisticated back-of-the-envelope calculations before beginning an econometric analysis. First, why might convergence matter at all for the next 100 years of growth? Won't world growth over such a long period be dominated by growth of the frontier, presumably due to technology raising TFP? Second, is the gap between world and frontier income large enough that even slow convergence might add substantially to growth of the frontier?

What might be a reasonable range for the frontier growth rate? Frontier income has been growing at about 1.7 percent per year since 1970.[2] Of course, that growth rate might change in the future. Gordon (2016, p. 14) offers some estimates that suggest a plausible range for growth in the frontier. He reports that U.S. income grew at a 1.8 percent annual rate from 1870 to 1920 and 2.4 percent from 1920 through 1970. Notably, Gordon argues that future growth won't exceed the current 1.7 percent rate. Forecasting frontier growth over the next 100 years is a difficult and open question, as there might be further productivity slowdowns or there

[2] Data is from 1950 through 2014 in 2011 dollars, as given in the Penn World Tables 9.0 including updates through August 2016, 2016. See Feenstra et. al. (2015). "Frontier income" means U.S. GDP per capita throughout, while "rest of world" means all countries other than the U.S.; including other countries with income that is also at the frontier, but excluding the petro-states Brunei, Kuwait, United Arab Emirates, and Qatar, as well as Macao which has questionable GDP numbers and Curaçao and Sint Maarten (Dutch part) for are very small and for which the data only begins in 2005.

might be quickened growth due to increased automation. Taking 1.7 percent as a benchmark, we can ask if convergence is quantitatively important by asking whether growth forecasts are notably different from 1.7.

The *potential* for faster world growth through convergence is enormous, because most of the world is still very far from the frontier. Taking the United States as representative of the world frontier, as is customary in the post-War period, frontier income is now $51,600 while rest of world income is $12,600. If the rest of the world were to catch up to the frontier in a single, 20-year generation (miraculously), then the world growth rate would equal 9.2 percent. If convergence were complete in a less miraculous time frame, say, 50 or 100 years, then the world growth rate would have to be 4.7 or 3.2 percent, respectively. All these numbers are much larger than plausible rates of frontier growth. Thus convergence is potentially key.

The second back-of-the-envelope calculation looks at how in recent years convergence appears to have become more important than it was in the earlier postwar period. Figure 1 shows income in the rest of the world relative to the frontier on the left axis. At some point in the last decade of the twentieth century, relative world income began to grow notably more quickly than in the earlier postwar period. Importantly, between 1990 and 2014 relative income in China increased by almost a factor of four. Relative income in India doubled. Since China and India account for a third of the world's population, growth in these two countries alone can make a huge difference to long-run world income. The right-hand scale for Figure 1 shows my econometric estimates of the probability that the average, population-weighted, country has started to converge to the frontier, with separate estimates allowed pre- and post-1990. The methodology is discussed below. For now, Figure 1 illustrates that the changes in the estimated probability that countries are converging are largely consistent with the pattern one sees in the raw data for world income, and suggests that by the end of the sample about two-thirds of (population-weighted) countries had started on a convergent path.

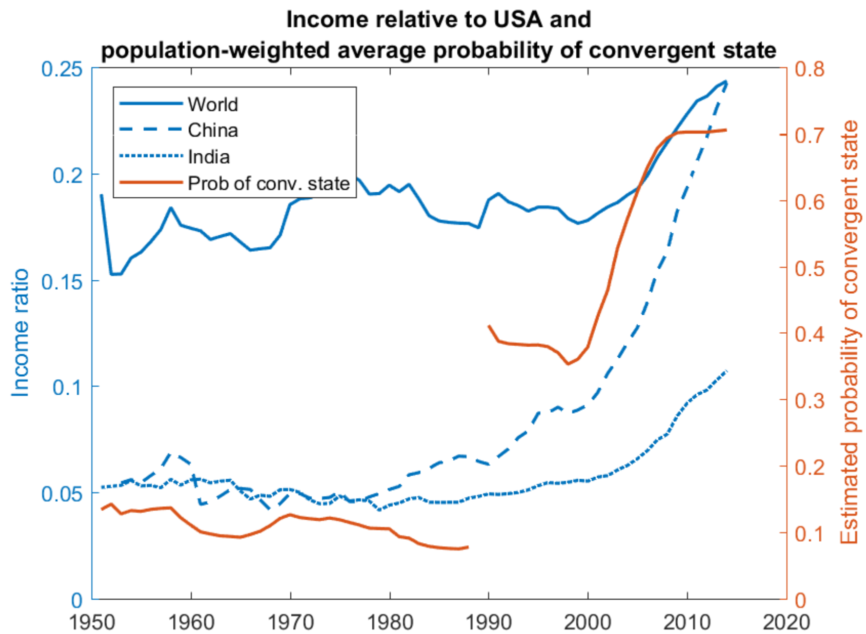

Figure 1

Continuing with simple calculations, suppose we think that convergence in recent decades might continue. Say there are two chances in three it will continue and one chance in three that the distance to the frontier will close no further, in other words that convergence will come to a complete halt. Suppose further that if a country is converging, it takes 50 years to close half the gap to the frontier. Together these assumptions imply that in 100 years there are two chances in three that the gap will be reduced to one quarter of its current level. That adds one percentage point above the frontier to the world growth rate.

Neither theory nor back-of-the-envelope calculations dictate that the next century should see complete convergence. Indeed, I estimate income in the rest of the world will rise from 24 percent of the frontier to 67 percent of the frontier level over the next hundred years. This suggests an overall growth rate of 2.7 percent, which is a percentage point higher than frontier growth. Predictions are shown in Figure 2. Without convergence, I estimate a hundred years of growth will put world income about sixty percent higher than today's U.S. income. With estimated convergence, world income will be nearly quadruple today's U.S. income. (U.S. income will be nearly six times today's income.) These are median point estimates; fiducial intervals are presented below.

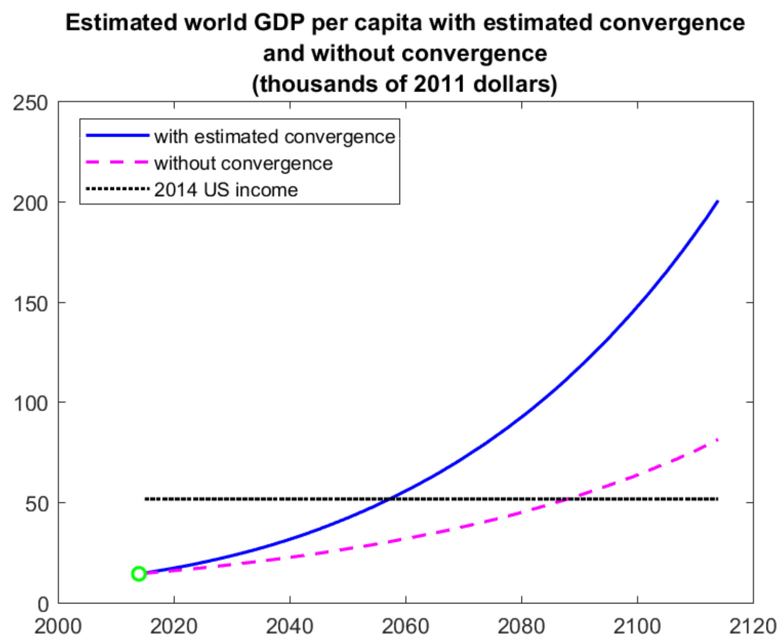

Figure 2

In the next section, I present an illustrative model to nail down concepts and compare the model to Lucas' (2000). I then discuss the data and the estimation technique, including the role of the identifying restrictions. I then make estimates of the state of convergence for all the countries in the world and use the estimates to project income for each country for the next century.

I. Models and related literature

Many models of economic growth, including the one presented below, explicitly suggest that at some point a country is likely to switch growth paths. As Jerzmanowski (2006) writes, “growth experiences differ over time within a country almost as much as they differ among countries.” Hausmann et. al. (2005) report over 80 instances of growth accelerations between 1950 and 2000. Jones and Olken (2008) document that growth can stop as well as take off. [3]

Projections of long-run world growth are often done using “scenarios” or calibration, unlike the purely econometric methods used here. The work closest in spirit to this paper is Lucas (2001). Lucas (Figure 3) forecasts a growth rate starting around 3 percent and declining to about 2.2 percent over the period in which I forecast. (Lucas assumes slightly higher frontier growth.) Other forecasts often project out factor inputs and then run these inputs through a production function. Examples include Johansson et. al. (2012) and Dellink et. al. (forthcoming). The econometric forecast closest to the one here is Zimmer et. al. (2016), which estimates convergence rates for each country in the world based on historical convergence rates. Using Bayesian techniques somewhat similar to those used here, but assuming that countries never change convergence paths, Zimmer et. al. (2016) estimate that for over half the countries the rate of convergence is zero. As a result Zimmer projects world income to grow at little more than the rate of frontier growth. Jerzmanowski (2006) offers a Markov-switching model of growth, focusing on the role played by government anti-diversion policies. Jerzmanowski allows four states of growth (for individual countries rather than relative to the frontier) and the data

[3] Kraay and McKenzie (2014, p. 128) argue that no-growth states are rare in the modern era, writing “stagnant incomes…over long periods is rare in practice, with the typical poor country growing at least as fast as the global average over the last 60 years.”

set ends in 1996. The Bayesian technique used here helps separate long-run from business cycle dynamics in a way that cannot be done using the maximum-likelihood estimation in Jerzmanowski. These differences notwithstanding, Jerzmanowski's results suggest that the Markov-switching approach is a fruitful way to identify growth regimes.

The paper with the closest theoretical model to the one used here is Lucas (2000). Lucas, focusing on the period since 1800, has countries either in no growth status or in convergent status. A given country has no growth until it begins convergence, which happens probabilistically. The estimates here, based on the modern era, look at growth relative to the frontier rate without or with convergence. The mechanics are the same when there is convergence, but the alternative is no growth in Lucas or conditional convergence here.

I present here a simple model to illustrate the assumed mechanics. Let Y be per capita GDP. A country has two production sectors. Each country has one unit of resource per capita, divided between the developing sector and the modern sector, $X_D + X_M = 1, 0 \leq X_D, X_M \leq 1$. Both sectors are exposed to the world technology frontier, A, but the developing sector makes inefficient use of the technology. Production is $Y_D = \beta A X_D, 0 < \beta < 1$ and $Y_M = AX_M$. For countries at the world technology frontier we have $X_M = 1$,and income $Y^* = A$. For other countries, $Y = Y_D + Y_M = \beta A X_D + AX_m$. The ratio of a country's income compared to the world frontier is $Y/Y^* = \beta X_D + X_M$, so the growth rate of the ratio is

$$[Y\dot{/}Y^*] = (1-\beta)\dot{X}_M \tag{1}$$

There is a continuum of agents endowed with time that can be turned effortlessly into one unit of X_D in aggregate. Alternatively, agents can engage in costly effort to create X_M. The cost of engaging in the modern sector is decreasing in the size of the existing modern sector. The dollar-equivalent utility cost of creating modern input for agent j is $\lambda_j - \ell X_M$, where λ_j—the idiosyncratic part of the cost—is distributed across the population with density $f(\lambda)$. An agent is indifferent between engaging in the two sectors if $\beta A = A - (\lambda_j - \ell X_M)$, so the marginal agent has cost

$$\lambda^* = A(1-\beta) + \ell X_M \tag{2}$$

Allocation of inputs is determined by

$$X_M = \int_{-\infty}^{\lambda^*} f(\lambda) d\lambda \tag{3}$$

If the cost of the marginal agent is below the lower support of $f(\lambda)$, $X_M = 0$ and the economy does not converge. If λ^* is greater than the upper support $f(\lambda)$, $X_M = 1$ and the economy is on the frontier. In either case, the gap between country income and frontier income is constant. In between, the gap closes according to equation (1). If one assumes for convenience that $f(\lambda)$ is uniform with density c inside the support, then from equation (3) we have $\dot{X}_M = c\dot{\lambda}^*$. From equation (2) we have $\dot{\lambda}^* = \dot{A}(1-\beta) + \ell \dot{X}_M$. Together with equation (1) we have that for interior solutions the gap closes at rate

$$[Y/\dot{Y}^*] = (1-\beta)\frac{c(1-\beta)}{1-c\ell}\dot{A} \tag{4}$$

I close the model by assuming that log frontier income, y^*, follows a random walk with growth rate $\mu > 0$. Positive frontier growth means that A eventually becomes high enough that that all economies eventually start to converge because λ^* will pass the lower support for λ. Thus, as with Lucas, the *very* long-run prediction is that all countries will converge—although not necessarily in a mere 100 years.

In order to bring the model to the data, I simplify by assuming that all countries have the same convergence rate when they have an interior solution; approximating equation (4) with logs, $y = \log(Y)$, and move it into discrete time; and add the possibility of random shocks.[4] In principle a large enough random shock could cause a country to revert from a convergent to a non-convergent path. More importantly, countries sometimes undergo growth disasters in which they diverge from the frontier for extended periods (see Pritchett (1997)). Thus I extend the model to allow three discrete states: non-convergent (n), convergent (c), and divergent (d). The model to be estimated becomes

$$\begin{aligned} y_{i,t} - y_t^* = I_{S_{i,t}=n} \times \left[y_{i,t-1} - y_{t-1}^*\right] + I_{S_{i,t}=c} \times \rho_c \times \left[y_{i,t-1} - y_{t-1}^*\right] \\ + I_{S_{i,t}=d} \times \rho_d \times \left[y_{i,t-1} - y_{t-1}^*\right] + \varepsilon_{i,t} \end{aligned} \tag{5}$$

$$y_t^* = \mu + y_{t-1}^* + \varepsilon_t^* \tag{6}$$

where countries are indexed by $i = 1, \dots, n$. $S_{i,t} \in \{n, c, d\}$ during the convergence period, so the log ratio between a country's income and the frontier closes at an expected rate of $1 - \rho_c$ percent, where ρ_c summarizes the information in $(1-\beta)\frac{c(1-\beta)}{1-c\ell}$. Outside the convergence period, if $S_{i,t} = n$, the gap is unchanged except for random fluctuations and if $S_{i,t} = d$ income falls relative to the frontier. In this specification a country that has caught up to the frontier is expected to stay there in both states n and c, except for random fluctuations. For such countries the state is not identified, but this is of no consequence because the predictions are the same. The model accommodates a country which has an idiosyncratic steady-state that

[4] Assuming a common value for ρ surely oversimplifies the world. Note, however, that some version of such an assumption is necessary as many countries have few observations with $S_{it} = c$. Essentially, ρ must be inferred from countries that are on the convergent path but not yet converged. In principle, some heterogeneity could be added by making ρ a function of observables. Doing so might be interesting for its own sake, but doing so would then add the difficulty of forecasting out those conditioning observables for 100 years in order to know future ρ.

leaves it permanently below the frontier by allowing the country to begin in the convergent state but eventually switch to the non-convergent state as it approaches its permanent deviation from the frontier. Finally, note that equations (5) and (6) are very close to equations (1) and (2) in Lucas (2000).

To complete the empirical specification, I take $S_{i,t}$ to follow a discrete first-order Markov process. (Jones (1997) also uses a Markov specification in a somewhat different way to study the distribution of world income. Jones emphasizes the importance of convergence in determining the world income distribution. Durlauf and Johnson (1995) divide countries into different growth regimes, although not using a Markov-switching model.)

II. Data

The data used here is from the Penn World Tables (Feenstra et. al. (2015)) version 9.0. GDP per capita is measured in 2011, PPP converted, U.S. dollars and is calculated according to the PWT variables $Y = RGDPNA/POP$. The data measures income for 174 countries annually with various starting dates beginning in 1950 and all ending in 2014, for a total of 9,129 observations. Table 1 provides descriptive statistics of the underlying data, showing both population-weighted and unweighted descriptive statistics. [5] Observations start in various years; the weighted mean first observation comes in 1955. The shortest observation period is 25 years; the longest is 65. The ratio of average income to frontier income declined in the data set from individual country starting dates through 1953, but this largely reflects later entry into the data set of poorer, smaller countries. Note that the population-weighted ratio of country income to the frontier closed moderately between the first-observation for a country and 2014.

[5] Weighting is done throughout using 2014 population weights.

	Mean (weighted by 2014 population)	Mean (unweighted)	Min	Max
Starting year	1955	1963	1950	1990
Average growth rate, rest of world	2.71%	1.97%	-1.71%	7.14%
Average growth rate, U.S.	1.99%			
2014 income, rest of world	$12,571	$16,594	$570	$82,297
2014 income, USA	$51,621			
Ratio of income to US income, first observed year	0.13	0.40	0.01	3.68
Ratio of income to US income, 2014	0.24	0.32	0.01	1.59
Source: Penn World Tables 9.0. Data in 2011, PPP-adjusted U.S. dollars				

Table 1

PWT provides an unbalanced panel. There is likely some selection bias in terms of when coverage begins for a particular country. However, 89 percent of the world's population, by 2014 population weights, is covered by 1960. Figure 3 shows population coverage as well as the population-weighted, observed gap from the frontier, $Y_{i,t}/Y_t^*$. The average country experienced a slowly closing gap except for the first four years of data, which covered a much smaller fraction of the population than did the remainder of the sample.

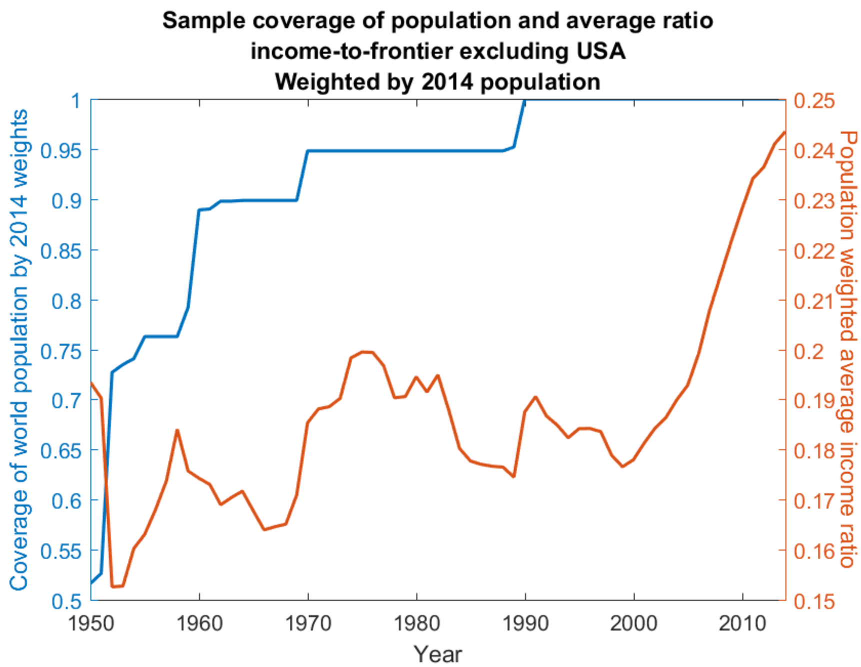

Figure 3

The theoretical model suggests positive growth at or above the rate of frontier growth, except perhaps for random fluctuations. Postwar data is not entirely consistent with the theory, but the deviations represent a very small part of world population. Mean growth was negative for 17 countries, although significantly negative with a one-tailed, five percent test for only one. The model also expects a country's growth to be equal to or higher than frontier growth, again with the exception of random shocks. Quite a few countries, 89, had mean growth lower than the frontier, but only 21 had growth significantly below the frontier on a one-tailed, five percent test of the difference between country and U.S. mean growth rates. Total 2014 GDP of the countries with growth significantly below the mean was 1.4 percent of world GDP, with almost half of that accounted for by South Africa.

III. Estimation

Estimation follows a fairly standard Bayesian algorithm, which is described first. I then describe the important identifying restrictions. A more detailed description of the algorithm, the priors, and the posteriors is given in the Appendix. The model consists of 174 equations for the income gap plus a random walk equation for frontier income, as specified in equations (5) and (6).

The frontier equation is estimated assuming normal errors and independent normal-gamma priors. The priors are diffuse, so that the posteriors mimic the standard frequentist distribution for a sample mean. μ and $h^* \equiv 1/\sigma_{\varepsilon^*}^2$ are estimated by Gibbs sampling with 100,000 samples retained after discarding 10,000 samples. The significant issue is that the U.S. underwent a growth slowdown somewhere around 1970; Gordon (2016) uses the date 1970 and Perron (1989) prefers 1973. Since projections for the future should reflect current growth, I estimate μ beginning in 1970.

The individual country equations in equation (5) are estimated by a Gibbs sampler, again with 100,000 samples retained after discarding 10,000 samples, iterating over conditional draws for the state vector S_i for each country, the precision $h_i \equiv 1/\sigma_{\varepsilon_i}^2$ for each country, and common values for ρ_c and ρ_d, and for probabilities of switching between states. Holding these parameters in common permits identification for countries in which a particular state is rarely observed, meaning there would be no data to permit idiosyncratic identification. The switching probabilities are estimated separately for the two subperiods; other estimates are common across the entire sample. The break at 1990 is chosen as a year in which there is arguably a change in the raw data, as seen in Figure 1, with some attention to leaving an adequate sample length in the later sub-sample. In other words, the results should not be interpreted to say something special changed precisely in 1990.

Conditional on S, I assume a uniform prior for ρ and an independent gamma prior for h_i. The priors for the precision are diffuse, with the prior mean for each country set to the unconditional variance of $y_{i,t} - y_t^*$ and the prior standard deviations set to four times the prior means.

The advantage of the Bayesian technique is that it allows for imposing identifying restrictions and for being explicit about what is being imposed. (Gibbs sampling also allows for considerable computational simplicity.) We want state estimates that are consistent with the economic idea that at a point in time economies either are or are not on a convergent path, and that once on the path generally stay there. Avoiding picking up higher-frequency business cycle coordination requires keeping $\rho_c < 1$, but not too far from 1.0. I do this with an informative uniform distribution that bounds ρ_c away from 1.0. At the lower bound of the prior, a country would close half the income gap from the frontier (starting at the average gap) in an implausibly fast 21 years. At the upper bound, the half-life of the average gap is 54 years. The priors for ρ_d are then set to be symmetric around 1.0 with the priors for ρ_c. The half-life to close the average gap at the posterior mean is 45 years.

The first-order Markov process for the state-switching model is controlled by the probabilities of remaining in the respective states, which are represented as separate Dirichlet processes for each subperiod. Presumably, switches out of convergence are also fairly rare. Picking a prior that enforces this helps identify low-frequency rather than business cycle behavior. I pick prior probabilities for state switches that place heavy weight on the probability of remaining in a given state. The posteriors reinforce this for states n and c, but not for the divergent state before 1990 where switches into the non-convergent state are quite probable. The steady-state probabilities evaluated at the posterior means for the early subperiod are $\{p_n = 0.834, p_c = 0.094, p_d = .073\}$. For the post-1990 subperiod, the posterior steady-state probabilities are $\{p_n = 0.175, p_c = 0.764, p_d = .062\}$. The 95 percent highest posterior density interval for the steady-state probability of being in the convergent state is $p_c \in (0.003, 0.269)$ in the early subperiod and $p_c \in (0.547, 0.973)$ in the later subperiod In other words, the posterior estimates suggest that the world has changed from a situation in which countries were largely not converging to a situation in which countries largely are converging.

The Gibbs sampler generates a posterior distribution for the convergence state, S, for each country in each year. Draws for each country are done separately, and draws depend on the entire history of the respective country's ratio relative to the frontier, as well as the Markov probabilities for the relevant subperiod. The mean of the posterior $S_{it} = j$ gives the probability that country i is in state j in year t. The right-hand scale of Figure 1 shows the increase over the sample period in the estimated population-weighted probability in being in the convergent state. The estimates suggest that convergence was unlikely in the early subperiod but quite likely in the later subperiod, particularly after the turn of the century.

Figure 4 shows the probabilities for all three states and also shows state estimates when the model is estimated over the entire sample (dashed line). The average probability of being on a convergent path rose from 11 percent in 1960 to 71 percent in 2014. Based on the posterior estimates, the probability that Korea was in a convergent state in 2014 was 94 percent. The probability for China was 98 percent and the probability for India was 96 percent. But the probability that Pakistan had moved into a convergent state was under four percent.

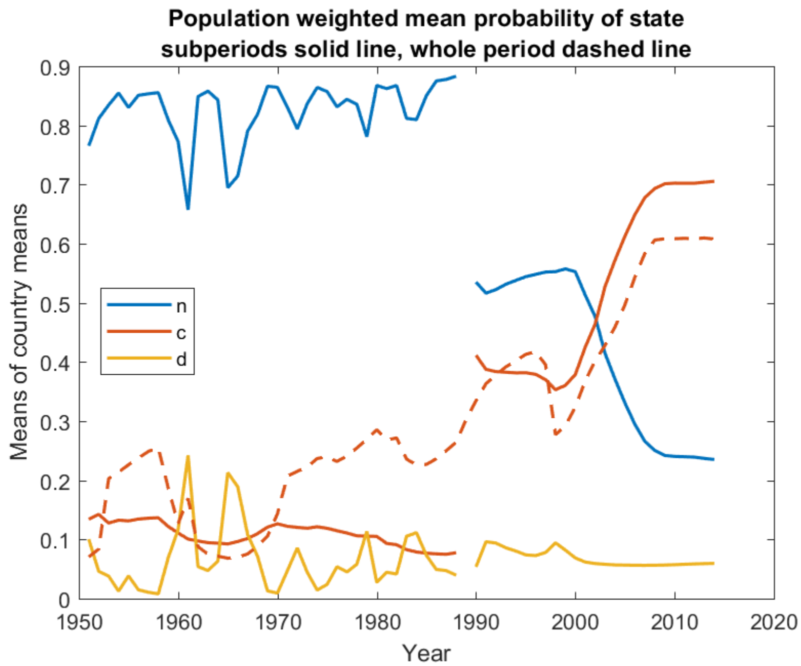

Figure 4

The apparent break in the state estimates is an artifact of estimating transition probabilities separately for the two subperiods. The whole sample estimate shows essentially the same pattern, but with the change between subperiods smoothed over. Since splitting the sample makes only a modest difference to convergence estimates, why split it? The answer is that while the in-sample estimates are largely driven by the data, forecasts at long horizons are significantly influenced by the steady-state probabilities. As shown above, these steady-state probabilities are estimated to have changed significantly in more recent years. Forecasts made using a whole sample estimate are qualitatively similar, but quantitatively different. The 100-year forecast growth rate based on the whole sample estimate is 2.3 rather than 2.7 percent.

As suggested above, what really matters is that several large countries are estimated to be on a convergent path. Figure 5 provides a bubble plot where the horizontal axis gives the income ratio to the frontier, the vertical axis gives the estimated probability of being in the convergent state, and the size of the bubble is proportional to a country's population. The diamonds give positions in 1970 and the circles give positions in 2014. The probabilities of convergence for the well-off countries toward the right of the chart do not matter terribly much simply because these countries are already close to the frontier. What matters a great deal is the large countries which appear to have moved onto a convergent path and whose income, while greatly increased, still has room to move significantly close to the frontier over the next 100 years. Note specifically the two large superimposed diamonds at the lower left which by the end of the sample period had moved to the circles at top and further to the right on the chart: China and India.

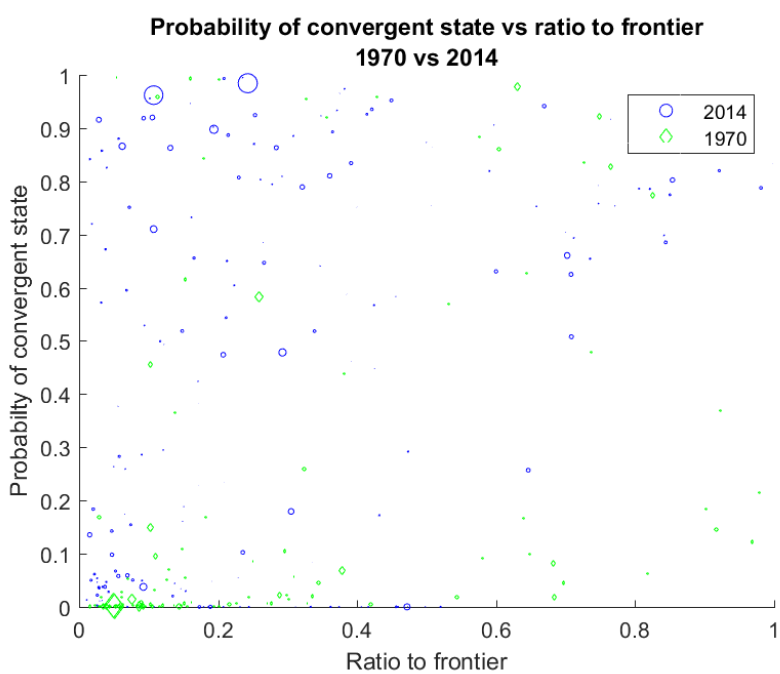

Figure 5

IV. Projections

Predictions over the next 100 years are made by taking 10,000 draws from the posterior distributions and then projecting the frontier and the gap from the frontier for each country. I begin with a draw for μ and h^*. The frontier is then forecast by taking 10,000 100-long normal draws for ε^* and running out the random walk with drift in equation (6). Next, draws are taken for the remaining parameters and the latent $S_{i,2014}$. With respect to the latter, note that while each state draw is either zero or one, the expected value of each draw is the posterior mean for each country. The states are then forecast out according to a first-order Markov process governed by the draw from the posterior for the 1990-2014 transition matrix. Finally, 10,000 100-long normal draws are taken for $\varepsilon_{i,t}$ and income is projected for the country using equation (5).

As shown in Figure 2, the median forecast for world income in 100 years is $200,000. This compares to a forecast frontier income of $293,000. In contrast, absent further convergence world income would be predicted to rise only to $81,000. Of course, each of these numbers would be changed if the average growth rate of the frontier over the next 100 years turns out to be significantly higher or lower than the average growth rate over the last 40.

A large part of the forecast of movement toward the frontier, 28 percent, arises from the forecasts for two countries: China and India. This may be unsurprising, as China and India have 37 percent of the world's population. (Indonesia is the only other country to account for as much as two percent of the result.) The estimated model does not make use of idiosyncratic information about likely future growth prospects for individual countries, above and beyond the information encapsulated in the historical growth record. The fact that nearly all countries play small roles in the overall world forecast suggests mis-forecasts for individual countries are unlikely to have much effect on the final result so long as one believes that China and India are likely to continue to converge.

As seen in Figure 2—which gives the median forecast—world income is projected to grow substantially over the next 100 years. Since the forecasts are based on how much countries have converged to the frontier, it is useful to look at projections for the gap from the frontier and the probability of being in a convergent state. Forecasts are shown in Figure 6. The 100-year forecast is that the ratio of income to the frontier will have risen from 24 to 67 percent. The population-weighted probability of being in a convergent state is forecast to rise slightly. In other words, the growth forecast largely reflects growth in countries that have already started the convergence process rather than a forecast future increase in the fraction converging.

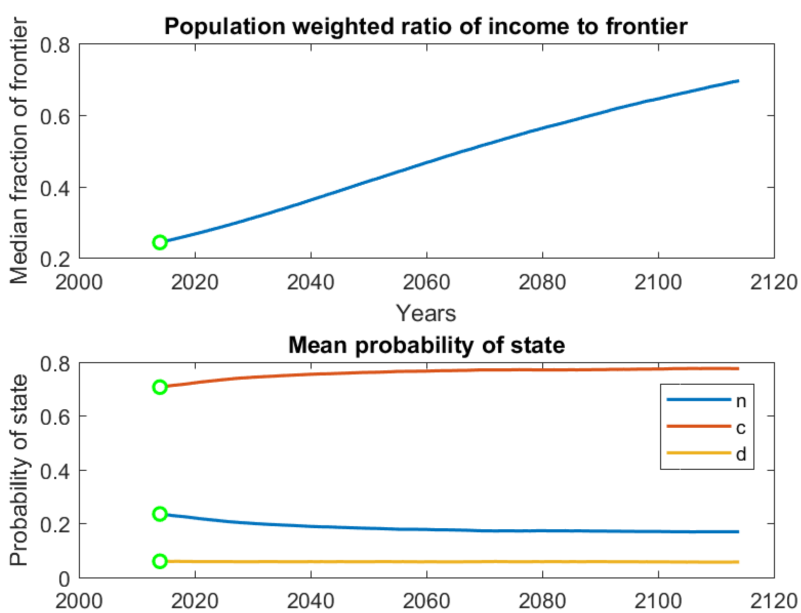

Figure 6

Uncertainty about future convergence comes from two sources. The first source of uncertainty is due to uncertainty about future shocks being in the convergent state and to future additive errors in equations (5) and (6). The second source of uncertainty is due to estimation of the parameters, i.e. posterior spread. This second source can be examined by comparing projections based on median posterior values for the parameters, while retaining posterior uncertainty about $S_{i,2014}$.

Figure 7 gives median projections and 90 percent highest posterior density intervals both with and without parameter uncertainty and with and without shocks to frontier growth. Most uncertainty arises from real shocks to equation (5) and to state uncertainty; parameter estimation uncertainty plays only a small role.

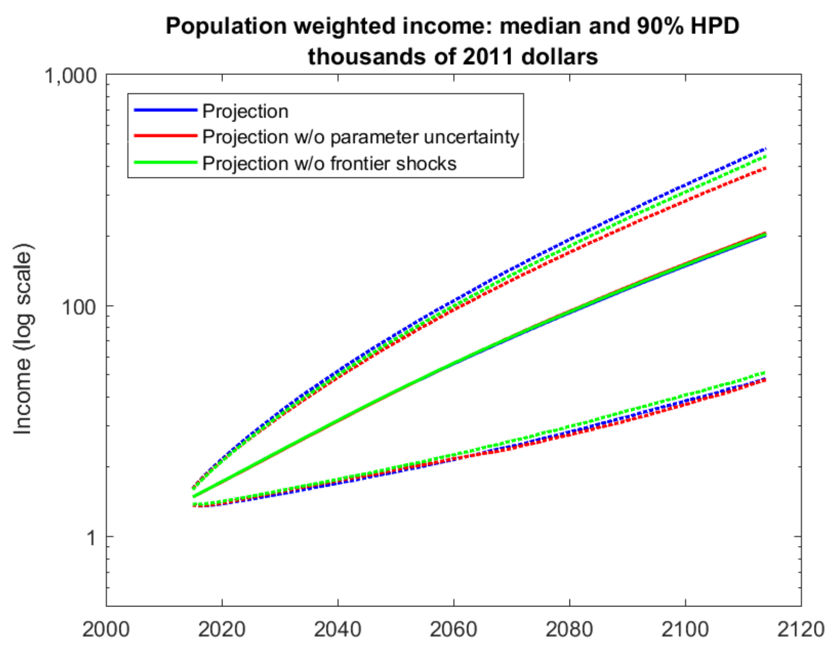

Figure 7

Lucas (2000) wrote "I think the restoration of inter-society income equality will be one of the major economic events of the century to come."[6] While the principal topic of interest here is forecasting future world income levels, the model also generates forecasts of income distribution. Figure 8 graphs the fraction of the world's population living at or below a given income level in 2014 and one hundred years later, using median country forecasts. In 100 years, most of the world's population will be in countries with far higher income levels than today's U.S. income. Inter-country income equality will also have increased significantly. The 2014 Gini coefficient is estimated to be 0.58. 100 years later the Gini coefficient is estimated to be 0.28.

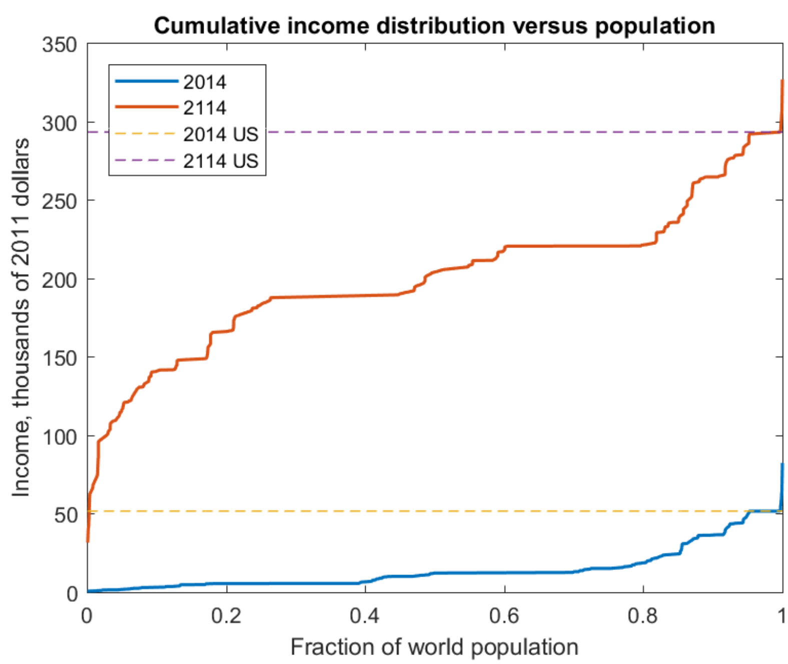

Figure 8

Conclusion

In 100 years, most of the world's population will live in countries with higher income than the United States today. And the majority of the world's population will have a much higher

[6] For a differing view, see Pritchett (1997).

income than the United States today. The typical standard of living will be much, much improved, and inter-country inequality will be lower. This forecast is based on the assumption that the pattern of convergence observed in the last quarter century will not be reversed.

A good forecast for 100 years hence is that income will be roughly quadruple today's U.S. income and that most of the world's population will be living in countries with high incomes by today's standard. Such growth reflects a great increase in the number of countries which have begun to converge to the world income frontier in recent years. The forecast is notably higher than is found when based on average postwar experience rather than focusing on more recent years. Both the raw data and the econometric estimates suggest that the world has changed in a way that makes convergence much more important than it may have appeared to be based on research that relied on earlier data.

Appendix

Details of the Bayesian estimation procedure are provided here. Estimation is by Gibbs sampling, with 100,000 draws retained after a 10,000 draw burn-in. The parameters to be estimated are μ, ρ_c, ρ_n, the 3×3 state probability transition matrix for each subperiod, the precision, h_i, for the shocks for each country in equation (5), and the precision, h^*, for the shocks to the frontier in equation (6). The states, S_{it}, are estimated as latent variables. Diffuse priors are used for μ, h^*, and h_i. The priors for ρ_c, ρ_n, and the transition matrix are chosen so as to be informative in order to identify low frequency rather than business cycle states. Since the states are created as a standard form of data augmentation, no priors are specified for S_{it}. Both ε_{it} and ε_t^* are assumed to be independent normal.

The frontier growth equation is estimated by a standard Bayesian regression assuming independent normal-gamma priors.[7] As discussed in the main text, the U.S. underwent a growth slowdown beginning around 1970. For this reason, equation (6) is estimated using data from 1970 onwards. Priors and posteriors for μ and h^* are given in Table A1. As a practical matter, the posterior mean and median for μ match the sample mean growth to two digits.

Parameter	Conditional prior	Posterior median	95 Percent HPD	Posterior mean	Posterior standard deviation
μ	$N(0, .0016) \times \Gamma((4.1 \times 10^{-4})^{-1}, 1)$	0.017	$< 0.011, 0.023 >$	0.017	0.0031
h^*		2388	$< 1462, 3436 >$	2,424	510

Table A1

[7] There are several parameterizations of the gamma distribution in the literature. I use $f_\Gamma(z|m,\nu) \propto z^{\frac{\nu-2}{\nu}} \exp\left(-\frac{z\nu}{2m}\right)$.

The values of ρ_c, ρ_d, and h_i are estimated using standard Bayesian regressions assuming uniform priors for ρ_c and ρ_d and an independent gamma prior for h_i. Country i at observation t contributes to the estimation of ρ_c conditional on the draw $S_{it} = c$ and contributes to the estimation of ρ_d conditional on $S_{it} = d$. This gives truncated normal posteriors for ρ_c and ρ_d. Somewhat surprisingly, the draws from the truncated normal sometimes run into numerical difficulties, which are solved by using the algorithm given in Botev (2016) as well as code the author makes publicly available.

Each country's data is weighted according to its 2014 population share, although weighting the regression makes very little difference. The prior for ρ_c is chosen, as described in the main text, to allow a plausible range of speeds of convergence. The prior for ρ_d is chosen to be symmetric with the prior for ρ_c around 1.0. The posterior for ρ_c shows considerable pile up close to the upper bound—approximately 34 percent of the simulation draws—indicating the prior restriction does play an important role in identification. The posterior for ρ_d has a small pile-up—not quite 10 percent—near the upper bound. Table A2 gives information for the priors and posteriors for ρ_c and ρ_d as well as for the average precision parameter. The prior for the precision for individual countries is $\Gamma\left(\mathrm{var}(y_{it} - y^*), \frac{1}{8}\right)$, which is quite diffuse.

Parameter	Conditional prior	Posterior median	95 Percent HPD	Posterior mean	Posterior standard deviation
ρ_c	$U(0.95,0.98)$	0.979	$< 0.963, 0.980 >$	0.978	0.005
ρ_d	$U(1.02,1.05)$	1.036	$< 1.022, 1.05 >$	1.037	0.0087
$\bar{h}$	$\Gamma(51.2, 0.125)$	536	$< 29.6, 1{,}613 >$	409	118

Table A2

The prior for the nine parameter (less three adding up constraints) transition probability matrix is chosen to place a strong weight on a country remaining in a given state, particularly that once a country begins to converge it is likely to continue to do so as suggested by the theoretical model. The priors for states n and d are symmetric. The prior for state c is slightly more persistent, and makes it slightly less likely to transit into state d than into state n. The Dirichlet distribution provides conjugate priors. In order to ease interpretation, prior means are shown first and prior standard deviations are shown in parentheses. The parameters for the Dirichlet priors are given in square brackets in Table A3.

From \ To	n	c	d
n	0.965 (0.028) [39.3]	0.0175 (0.020) [0.714]	0.0175 (0.020) [0.714]
c	0.018 (0.022) [0.714]	0.973 (0.025) [39.3]	0.0088 (0.015) [0.357]
d	0.0175 (0.070) [0.020]	0.0175 (0.070) [0.020]	0.965 (0.005) [0.028]

Transition Matrix Priors

Table A3

The conjugate posterior is Dirichlet with parameters equal to the prior parameters plus the number of switches. Because there are just under 9,000 switches observed—on average 1,000 per element of the transition matrix—the data should dominate the prior in computing the posterior. In calculating the number of switches, I weight according to 2014 population. Specifically, if country i has S_i^{jk} switches from state j to state k and if country i has a fraction

w_i of the 2014 world population, and the prior Dirichlet parameter is p_{jk}, then the posterior Dirichlet parameter is $p_{jk} + \sum_i \left(S_i^{jk} \cdot w_i\right) \Big/ \frac{\sum_{jk}\left(\sum_i S_i^{jk} \cdot w_i\right)}{\sum_{jk}\left(\sum_i S_i^{jk}\right)}$.

Posterior means, medians (in braces), and 95 percent highest posterior density regions for the transition matrices are given in Table A4 and Table A5. The posteriors for states n and c are relatively close to the priors; the posterior for state d in the early subperiod is less so.

Given a transition matrix, we can compute steady-state probability distributions. It is useful to remember that long-run predictions for countries close to the frontier are essentially the same whether the countries end up in state n or state c. The steady-state probabilities and associated HPD's are given in the main text.

To From	n	c	d
n	0.958 {0.961} <0.926, 0.986>	0.002 {0.001} <0, 0.009>	0.039 {0.037} <0.012, 0.072>
c	0.012 {0.005} <0, 0.046>	0.966 {0.973} <0.913, 1.0>	0.023 {0.014} <0, 0.074>
d	0.458 {0.460} <0.289, 0.630>	0.017 {0.007} <0, 0.063>	0.525 {0.525} <0.364, 0.682>

Transition matrix posteriors—early subperiod

Table A4

To / From	n	c	d
n	0.957 {0.957} <0.931, 0.978>	0.042 {0.042} <0.019, 0.066>	0.001 {0.0} <0, 0.002>
c	0.007 {0.003} <0, 0.025>	0.990 {0.992} <0.973, 1.0>	0.0040 {0.003} <0., 0.009>
d	0.004 {0.004} <0.002, 0.082>	0.006 {0.003} <0, 0.023>	0.954 {0.953} <0.912, 0.996>

Transition matrix posteriors—later subperiod

Table A5

The states S_{it} are treated as latent variables. Conditional state draws are made using a multi-move algorithm based on Carter and Kohn (1994). (See also Chib (1996), and Kim and Nelson (1999).) The standard algorithm is modified to restrict a draw of the divergent state to cases where the country's income is below the frontier in a given year. This is consistent with the model and avoids drawing a small number of explosive paths (primarily for Norway). In other words, when $y_{it} > y_t^*$ a random draw of d is recoded to n. Posterior means for each state are shown in Figure 4 in the main text.

The countries with notable posterior estimates of being in a divergent state at the end of the sample include the Central African Republic (48%), Greece (43%), Cayman Islands (41%), Italy (38%), and Syria (30%). No other countries are as high as 20 percent at the end of the sample.

References

Acemoglu, Daron. 2009. *Modern Economic Growth,* Princeton University Press.

Barro, Robert J. and Xavier Sala-i-Martin. 1995. *Economic Growth,* McGraw-Hill.

Botev, Z.I. 2016. "The normal law under linear restrictions: simulation and estimation via minimax tilting," *Journal of the Royal Statistical Society: Series B (Statistical Methodology),* 79 (1), 125-148.

Carter, C.K. and R. Kohn. 1994. "On Gibbs Sampling for State Space Models," *Biometrika,* 81 (3): 541-553.

Chib, Siddhartha. 1996. "Calculating posterior distributions and modal estimates in Markov mixture models," *Journal of Econometrics,* 75 (1), 79-97.

Dellink, Rob, Jean Chateau, Elisa Lanzi, and Bertrand Magné. Forthcoming. "Long-term economic growth projections in the Shared Socioeconomic Pathways," *Global Environmental Change.*

Durlauf, Steven N., Paul A. Johnson, and Jonathan R.W. Temple. 2005. "Growth Econometrics," *Handbook of Economic Growth,* Vol. 1A, Philippe Aghion and Steven N. Durlauf, eds., Elsevier B.V.

Durlauf, Steven N. and Paul A. Johnson. 1995."Multiple Regimes and Cross-Country Growth Behavior," *Journal of Applied Econometrics,* 10 (4), 365-384.

Feenstra, Robert C., Robert Inklaar and Marcel P. Timmer. 2015. "The Next Generation of the Penn World Table," *American Economic Review*, 105 (10): 3150-3182.

Gordon, Robert J. 2016. *The Rise and Fall of American Growth,* Princeton University Press.

Hausmann, Ricardo, Lant Pritchett, and Dani Rodrik. 2005. "Growth Accelerations," *Journal of Economic Growth,* 10 (4), 303-329.

Jerzmanowski, Michal. 2006. "Empirics of hills, plateaus, mountains, and plains: A Markov-switching approach to growth," *Journal of Development Economics,* 81 (2): 357-385.

Johansson, Åsa, Yvan Guillemette, Fabrice Murtin, David Turner, Giuseppe Nicoletti, Christine de la Maisonneuve, Philip Bagnoli, Guillaume Bousquet, and Francesca Spinelli. 2012. "Looking to 2060: Long-term global growth prospects," *OECD Economic Policy Papers,* No. 03.

Jones, Benjamin F. and Benjamin A. Olken, 2008. "The Anatomy of Start-Stop Growth," *Review of Economics and Statistics,* 90 (3): 582-587.

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

________ and Dietrich Vollrath. 2013. *Introduction to Economic Growth, 3rd ed.,* W.W. Norton & Company.

Kim, Chang-Jin and Charles R. Nelson. 1999. *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications,* MIT Press, Cambridge, MA.

Kerekes, Monika. 2012. "Growth Miracles and failures in a Markov switching classification model of growth," *Journal of Development Economics,* 98 (2): 167-177.

Kraay, Aart and David McKenzie. 2014. "Do Poverty Traps Exist? Assessing the Evidence," *Journal of Economic Perspectives,* 28 (3), 127-148.

Lucas, Robert E. Jr. 2000. "Some Macroeconomics for the 21st Century," *Journal of Economic Perspectives,* 14 (1): 159-168.

Perron, Pierre. 1989. “The great crash, the oil price shock, and the unit root hypothesis,” *Econometrica* 57, 1361–1401.

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

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

Zimmer, Alec, Adrian E. Raftery, Dargan Frierson, and Richard Startz. 2016. “Less Than 2° C Warming by 2100 Unlikely,” Working paper, University of Washington.