Assessing Income Convergence with a Long-Run Forecasting Approach: Some New Results

Silva Lopes, Artur

6 February 2024

Assessing Income Convergence with a Long-Run Forecasting Approach: Some New Results *

Artur Silva Lopes †

This version: February 6, 2024

Abstract

Relying on low frequency econometric methods, a new simple procedure to assess international income convergence is introduced. It implements the long-run forecasting definition and discards short and medium-term information contents of the data as these may produce misleading evidence. Robustness to non-stationarities is achieved using first differences of (logged) per capita incomes.

Application to a selected sample of 90 different countries provides mixed but generally more positive evidence than most previous studies. Nevertheless, it casts many doubts on the inevitability of income convergence, at least in practically relevant time frames and as worldwide phenomenon.

Keywords: income convergence; growth empirics; low frequency; interval forecasting.

JEL codes: O47, C22, C53, N10.

*A preliminary and wrong version of this paper circulated with the title "Feels like going backwards: assessing income convergence with a long-run forecasting approach". I must apologize the readers of that version because a seemingly minor (sign) error radically changed many statistical results.

†Independent researcher. I am grateful to Richard Startz for helpful comments. I am also grateful to Prasada Rao for helpful suggestions. Obviously, the usual disclaimer applies. Email: artursl(at)meo.pt or arturslopes(at)proton.me.

1 Introduction

Economic growth obviously matters much, not only *per se* but also because "*it is the most powerful force for reducing global poverty and inequality*" (Milanovic, 2016, p. 232). The process through which it operates at this level is known as economic convergence, broadly meaning that per capita incomes of poor nations should get closer to that of the rich ones.

The elimination of these gaps is sometimes considered as an inevitable phenomenon, one that will certainly occur sooner or later; see, e. g., Lucas (2000) and Startz (2020). It is this view that supports its partial (Startz, 2020) or total (Müller, Stock and Watson, 2020) imposition on forecasting models, by means of restrictions on the parameters of those models, or by assuming that the same global growth factor affects all countries, respectively.

Empirical support to the hypothesis of economic convergence, however, is scant. Although a few exceptions prevent unanimity — e.g., Christopoulos and Leon-Ledesma (2011) and King and Ramlogan-Dobson (2014) —, the most common findings are usually far from supportive. Johnson and Papageorgiou (2020, p. 165) summarize a huge amount of evidence stating that "*there is a broad consensus of no evidence supporting absolute convergence in cross-country per capita incomes — that is poor countries do not seem to be unconditionally catching up with rich ones*".

This allows us to understand the reason why *The Economist* (2014) referred to the phenomena observed in the first decade of this century as an "*aberration*". Indeed, for a period of about 10 to 12 years, many poor countries, led by two of the most populous nations, China and India, were catching-up with rich ones at an unprecedented rapid speed, an episode that *The Economist* called "*The Great Catch Up*". But this process seems to have lost momentum about 10 years ago, and for many countries convergence appears to have become a mirage again.

The lesson to draw from this episode is quite clear: (in case it exists) convergence is really a long-run phenomenon, one that requires long periods of time to be assessed, one which does not conform with trends extrapolated on the basis of only a few years of observations. That is, short and medium-term variability of the data can be misleading about its long-run properties. The simple method that I propose in this paper avoids these shortcomings because it takes into account only low frequency, long-run variability of the data. This is achieved relying on low frequency econometric methods developed by Müller and Watson (2008, 2020 and mostly 2016, MW16 hereafter).

Some other features of the procedure that I propose appear to deserve some emphasis too:

a) it follows closely the well known convergence definition of Bernard and Durlauf (1996), which adopts a forward looking, long-run forecasting approach;

b) it simply extrapolates past real long-run trends without resorting to any particular formal model, thereby allowing to avoid making any specific assumptions;

c) it is very simply based in a comparison of two averages, the one that is needed to close the income gap and the maximum that is forecast to be attainable with a certain confidence level;

d) it has some robustness to non-stationarities, in particular to those that are due to breaks in the series, which is achieved through the application of the method to the first differences of the logged income gaps, not to the original (logged) series.

This last feature is particularly valuable because the occurrence of structural breaks along the path to the very distant future would be hardly surprising. And indeed, very distant horizons must sometimes be considered for lagging countries to catch-up with the technological leader. On the other hand, relying on c) precludes the possibility of forecasting the step by step future trajectories of countries' income gaps. The only object of estimation is the final position of such trajectories, corresponding to a certain predetermined forecasting horizon. Provided the only purpose is the estimation of these “arrival points”, the procedure can be easily and robustly employed in long-run forecasting in many different contexts, as shown in MW16. Since the sample ends in 2018, both the recent covid-19 pandemic and the wars in Ukraine and Palestine could not be considered.

In the context of time series empirical literature this paper provides mixed evidence about income convergence, generally more positive than is usual but, nevertheless, insufficient to grant the hypothesis the status of an unavoidable feature of modern economies. Furthermore, it must be emphasized from the outset that the analysis is not one of point estimation, investigating whether, in a number h of years, the incomes of countries will really attain or coincide with the level of the leader. Instead, the question that is answered here is: (allowing for some margin of uncertainty) is it very unlikely that in a time span of h years the predicted average rhythm for closing the gap with the U. S. is insufficient to close that gap? The question then amounts to investigate whether in the very long-run, defined here as a time horizon of $h = 42, 82$ and 100 years, each country will lie in a very wide “basin of attraction” around the position of the leader. Not whether it will really attain or even whether it gets close to that position. Sometimes I even mention that the country will likely enter the “orbit” of the leader, when the maximum attainable average equals or exceeds the average speed that is required to catch-up with the leader. Obviously, even entering that orbit is no guarantee that the country will eventually converge.

While not so bleak as in most previous studies, the evidence gathered here is sufficient to prevent the hypothesis to be considered as inevitable everywhere, sometime in a reasonably foreseeable future. At least, many doubts emerge that such a scenario will eventually materialize:

a) when the forecast horizon is 42 years “only”, barely more than half the population considered here (about 5.8 billion people) lives in countries forecast as attaining the mentioned “orbit”;

b) when the forecast horizon is extended to 2100 ($h = 82$), the number of “satellite" countries is obviously expected to increase but, even in the best case scenario, at least 2.1 billion people are still predicted to be left behind in living standards.

A further caveat to the positive results must be also made from the start: contrasting with usual time series approaches, convergence is taken here as the protected (null) hypothesis.

The remainder of this paper is structured as follows. In the next section the notion of convergence is revisited and its testing is briefly discussed. Section 3 describes the method proposed for assessing income convergence and section 4 discusses some preliminary issues of the empirical application. Section 5 presents the empirical results and section 6 analyses their robustness. Section 7 concludes the paper. In separate appendixes I first provide a brief methodological review, adding further details to the presentation in section 3, and finally I present empirical results for two less common levels of confidence.

2 Defining convergence

The most widely used formal definition of income convergence is due to Bernard and Durlauf (1996, BD), explicitly stated in a forecasting framework: "*countries i and j converge if the long-term forecasts of (log) per capita output for both countries are equal at a fixed time t*,

$$\lim_{k\to\infty} \mathrm{E}(y_{i,t+k} - y_{j,t+k}|\mathcal{F}_t) = 0,'' \tag{1}$$

$\mathcal{F}_t$ denoting the set of all information available at time t (including at least all historical information on both outputs). Recalling that the best forecast in mean squared error is given by the expectation conditional on the information set that is available at the time of the forecast[1], the definition requires that the very long-run

[1]Formally, in case one needs to make a forecast at time t, say of y_{t+h} (h denoting the forecast horizon), based on an information set that is available at that time, I_t, the best forecast in mean

optimal forecasts for both outputs must be the same, regardless of initial factor endowments.

But this forward looking and dynamic perspective has been transformed into a backward looking one, requiring only that past behaviour satisfies some mild conditions. A long-run one, true, but merely assessing whether there have been any past diverging signs. This is the case, for instance, when the discrepancy series is submitted to unit root tests. The initial long-run, real *future convergence* requirement was changed qualitatively; it was transformed into one of *past non-divergence*. Two countries *do not diverge* when the discrepancy in their (logged) income or output per capita is both stochastically and deterministically bounded.

Extrapolating this past limited gap behaviour into the future leads to the imposition that the discrepancy series must not be allowed to contain any deterministic linear trend, as argued in Lopes (2016, 2023)[2]. This is a possibility that is allowed in the usual (i.e., linear trending) notion of catching-up, but it is obvious that it would violate BD's condition: in the long-run the laggard would not only catch-up with the leader but it would overtake him as well; past the "meeting point", the positions of the two countries would invert and a symmetrical gap would appear and increase with time. (Linear trending) Catching-up can be useful only as a backward looking lens. On the other hand, assessing whether there was any (past) stochastic trend — which is the purpose of unit root tests —, is not sufficient to warrant that the initial BD condition holds. Only that it is not severely violated.

This change in nature is embedded in Pesaran's (2007) definition: *"countries i and j converge if for some finite positive constant C and a tolerance probability measure $\pi \geq 0$*

$$\Pr\{|y_{i,t+s} - y_{j,t+s}| < C|\mathcal{F}_t\} > \pi$$

at all horizons $s = 1, 2, \ldots, \infty$". This condition precludes the presence of deter-

squared error (MSE), the one that minimizes MSE, is the expectation of y_{t+h} conditional on I_t: $y^*_{t+h|t} = \mathrm{E}(y_{t+h}|I_t)$; see, e.g., Hamilton (1994), pp. 72-73.

[2]The ruling out of both unit roots and time trends by the convergence hypothesis was originally put forward by Quah (1993). Lopes (2016, 2023) emphasizes that linear time trends must not be allowed in unit root tests of the discrepancy series.

ministic and stochastic trends in the gap process and therefore corresponds more closely to the analysis provided by unit root tests. But, as acknowledged by Pesaran, it is only a condition for "*non-divergence*", much weaker than BD's. It was also proposed with a particular economic model in mind, where BD's condition implies that some rather unrealistic assumptions must hold. While in the unit root testing framework allowing level stationarity (instead of zero mean stationarity) is sufficient to achieve the relaxation of these assumptions, more generally, conformity only with a specific economic model is untenable. Moreover, although assessing whether equation (1) holds will always require resting on some extrapolation of past behaviour, using a specific economic model to do it should be avoided whenever possible.

3 The proposed method

The procedure that I propose does precisely this: it simply extrapolates past behaviour without resorting to any specific structural model. Moreover, it rests only on one assumption, one which is hardly debatable: that the first differenced gap series are covariance stationary, $I(0)$.[3] That is, rather than envisaging directly individual long-run forecasts of per capita GDP, the assessment is based on long-run prediction intervals for the growth (decline) rates of the countries' *gaps*, that is, for the differenced gaps of the logged series.

Focusing on the growth rates that are needed to eliminate the gap with respect to the leader allows me to circumvent non-stationarity issues. For instance, it is not necessary to know whether the gap series really contains an exact autoregressive unit root or whether a local-to-unity AR(1) model is more appropriate. Instead, the procedure begins by asking: what must be the average rate of negative growth or decrease of the gap for each country so that it vanishes in the long-term?

Formally, denoting with g_{kt} the gap or discrepancy series for country k ($g_{kt} =$

[3] $I(0)$ roughly means that the process is already stationary, i.e., that is does not need to be differenced to achieve stationarity. More precisely, it is weakly stationary with a positive long-run variance (e.g., Hansen, 2022, pp. 566-7).

$y_{kt} - y_{lt}$, l denoting the technological leader) and with $T+h$ the future year where the comparison will be made[4], i.e., where the catch-up process must be completed, the first step simply consists in determining $\overline{\Delta g^*_{kh}}$, the average rate of decrease that is needed for country k to close the gap in h years[5].

On the other hand, a prediction interval for the future (real) average rate of decrease of the gap is made with a pre-specified confidence level, $1-\alpha$, call it I^α_h, where I am dropping the country's index to simplify the notation. Hence, neglecting the possibility of overshooting, with a confidence level of $1-\alpha$ convergence in h years requires that $\overline{\Delta g^*_h} \in I^\alpha_h$;[6] otherwise, if $\overline{\Delta g^*_h}$ is too large, that is, if $\overline{\Delta g^*_h} \notin I^\alpha_h$ because $\overline{\Delta g^*_h}$ is larger than the upper endpoint of the interval, there is (forecasting) evidence that convergence is very unlikely to be attained in h years.

Now, since the purpose is long-run forecasting, the most appropriate approach is the one that is provided by low-frequency econometrics, as exposited in the work of Müller and Watson (2008, 2020) and mostly in Müller and Watson (2016, MW16), where the techniques are applied to forecasting. Therefore, attention is restricted to the low-frequency variation of the data series[7]. To extract this information from the series these are projected on a set of trigonometric functions of time, the regressors $\cos[(t-1/2)\pi j/T]$ for $j = 0, 1, \ldots q$, yielding the first q cosine transforms, $\widehat{\beta}^q_{LF}$ (denoted with $\mathbf{X}_{T,1:q}$ in UM16), q used to denote an upper limit to the frequencies that are considered, $q\pi/T$ (i.e., a lower limit to the corresponding periodicities, $2T/j$). Müller and Watson (2008) show that these projections provide a very good approximation to a close-to-ideal low-pass filter, i.e., one which retains only long-run variability and removes everything else from the data.

Based on a central limit theorem for both the OLS estimator for the coefficients

[4]As usual, T denotes the sample size and it represents as well the last observation in the sample, the time when the forecast is made; as is also usual, h denotes the forecast horizon.

[5]For instance, for Italy, in the last year of the sample the output gap for the U.S., in logged form, is -0.476; for an horizon of 82 years, the yearly average rate that is necessary to eliminate this gap is -0.0058, that is, an average rate of decrease of 0.58% per year.

[6]Overshooting means that $\overline{\Delta g^*_h} < \overline{\Delta g_L}$, with $\overline{\Delta g_L}$ denoting the lower limit of I^α_h.

[7]The relation between frequency and periodicity of the variations is an inverse one. Indeed, it is given by $p = 2\pi/\omega$, p denoting the period and ω the correspondent frequency, in radians.

of this regression and the average forecast value (for periods $T+1$ to $T+h$), MW16 show that an asymptotic two-sided $1-\alpha$ prediction interval for this average has endpoints given by

$$\overline{\Delta g_T} \pm t^q_{(1-\alpha/2)}(1+r^{-1})^{1/2}T^{-1/2}s_{LR}, \tag{2}$$

where $\overline{\Delta g_T}$ is the average rate of decrease in the sample data, $t^q_{(1-\alpha/2)}$ is the $(1-\alpha/2)$ quantile of the Student-t distribution with q degrees of freedom, $r = h/T$ and s_{LR} is an estimator of the long-run standard deviation, $s^2_{LR} = (T/q)\widehat{\beta}^{q\prime}_{LF}\widehat{\beta}^q_{LF}$ [8]. But since one needs a decision corresponding to a statistical test and since this must be one-sided — with non-convergence as the alternative —, the prediction interval must be one-sided as well. Therefore, there is evidence that there will be no convergence in the distant future whenever

$$\overline{\Delta g^*_h} \notin I^\alpha_h,$$

with

$$I^\alpha_h =]-\infty; \overline{\Delta g_T} + t^q_{(1-\alpha)}(1+r^{-1})^{1/2}T^{-1/2}s_{LR}], \tag{3}$$

$t^q_{(1-\alpha)}$ denoting the $(1-\alpha)$ quantile of the Student-t distribution with q degrees of freedom. That is, the country's effort to achieve convergence in h years is forecast that will unlikely succeed, falling short of the necessary gap decline whenever

$$\overline{\Delta g^*_h} > \overline{\Delta g_U},$$

$\overline{\Delta g_U}$ denoting the upper limit of the previous interval (which is country's specific)[9].

[8] This result is obtained using both a classical and a Bayesian approach, provided the prior is uninformative. Recall that the estimator of the long-run standard deviation allows for the presence of serial correlation in the series.

[9] In a separate appendix I provide further details of this method, as well as a few simple examples.

4 Preliminary issues

Before presenting the empirical results, some preliminary issues must be addressed. I describe them briefly below.

The data. The most adequate data set appears to be the one that was recently released by Bolt and van Zanden (2020), improving and updating the Maddison database and where, contrarily to the previous update, output per capita is represented by a single time series labeled GDP_{PC}. Generally, to employ the largest common sample size, the sample period begins in 1950 and ends in 2018, for a total of $T = 69$ observations. Several countries were excluded from consideration because their time series began only much later, implying a sample size that was too small. There are, however, several exceptions to this rule: a) for some Eastern European countries (Croatia, N. Macedonia, Montenegro, Serbia and Slovenia) the sample begins only in 1952; b) a major exception is Russia, whose time series begins only in 1960; a data series starting in 1950 (and extending to 2018) is available but for the former USSR and it will be considered in this study. Similarly, former Czechoslovakia will be considered as well, because the data series for the Czech Republic begins only in 1971 and the one for Slovakia only in 1985.

For comparison purposes with previous empirical results (based in a rather different method), I consider a core set of countries which is the same as in Lopes (2023). Besides the technological leader, the U.S.A. (simply U.S.), these are Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Hungary, Ireland, Israel, Italy, Japan, Netherlands, New Zealand, Norway, Poland, Portugal, South Korea, Spain, Sweden, Switzerland, and the U.K..

In this study, however, the set of countries is much enlarged, reaching a total of 89 countries besides the U.S. . With the exception of Ireland, Luxembourg, Norway, Singapore and Switzerland, whose per capita GDP in 2018 (in PPP USD of 2011) is higher than the one of the U.S., all these countries appear in table 1. A common restriction of previous time series studies — to confine the set of countries to those that are not far from their steady state — may now be relaxed.

The problem of poor power properties of statistical (unit root) tests originating from the consideration of economies in transition, pointed out in BD (p. 171), does not appear to be an issue now. Therefore, besides considering the so-called BRIC countries (Brazil, Russia, India and China), I have considered as well many countries that are very far from being considered as developed or high-income but that might attain convergence in a very long-run but foreseeable future. For these less developed countries the smaller per capita output in 2018 belongs to Ghana, with only 4267 USD in PPP units, representing only 7.7% of the one of the U.S. (and only about 1/20 or 5% of the one of Norway).

I have however excluded from consideration many economies highly dependent on exports of oil and natural gas, particularly those of the Middle East (as, e. g., Bahrein and Saudi Arabia), but also in Africa (say, Algeria, Egypt or Nigeria) and in South America (e.g., Ecuador and Venezuela). But I have considered countries as Jordan and Lebanon, also rich in these natural resources, but whose reserves started being exploited only very recently, in a process that I presume that has not affected the path observed until 2018. And of course, some major well known exceptions, such as Indonesia, Norway, Russia, the U.K. and the U. S., have been also included in the sample.

I have also excluded from consideration some low income countries (LICs), not only because for many of them data availability started only much later than 1950, but also because their inclusion would imply that too long forecast horizons would have to be considered.

In 2018, the total population considered here amounts to 6.137 billion people, i.e., 78.4% of the total population considered in Maddison's database. However, excluding population redundancies, that is, represented twice (as e.g., the population of former Yugoslavia), the sample of this study corresponds to 5.828 billion people, representing 77.7% of the total effective population.

Which horizon? That is, how long is the long-run? This is possibly the harder question. Taking into consideration the sample size, a decision made in MW16 ($h = 32$ for $T = 68$, that is, $h/T \approx 0.5$), and mostly the information provided by recent economic history, I arrived at a lower limit of $h = 42$, pushing the horizon to 2060.

This is much less ambitious than the horizon in Startz (2020), $h = 100$, but the method used here is even less structural, more "reduced form", than his. Nevertheless, his longest sample period, using data from the Penn World Tables (PWT) is also a bit smaller, beginning in 1950 but ending in 2014. The main advantage of my approach is that only long-run variability of the data is considered.

In forty years, between 1845 and 1885, the U.S. not only closed a gap of about 20% of the per capita GDP of the U.K. but it also advanced more 5%. And even discarding the case of Germany and the U.K. after the Second World War (only 12 years to close a gap larger than 65%, the *"Wirtschaftswunder"*, or "economic miracle"), there are several other cases even more successful than the previous change in the leader: between 1990 and 2006, i.e., in 16 years only, Ireland closed a gap of about 50% with the U.S.; in 30 years, between 1900 and 1930, Denmark has overcome a gap of about 35% of the output per capita of the U.K., and in only 25, between 1944 and 1969, starting from only less than 1/3 of its level, France has overcame the U.K. . As another example of a very successful process of catching-up in 40 years, between 1950 and 1990, Japan filled more than 60% of its gap with the U.S. (from 20% to 80.9%) and from only a bit more than 1/4 of the U.K.'s per capita GDP (27.6%), it attained a level significantly exceeding it, at 114%[10].

Forecast intervals deeper into the future, 82- and 100-year ahead, more realistic for many of the countries, will be also considered. Even more distant horizons could have been considered, but I believe that for convergence to be meaningful it must not be allowed to occur only in some undetermined future, particularly when technology appears to be so diffused. Moreover, for reasons that will become clear below, simply extending the forecasting horizon will produce a contraction effect in the width of the forecasting interval.

Which q? How low must be the lowest frequency? As can be readily observed

[10]Recently, the prime minister of India declared that by 2047 India will become a high income country, increasing its income per capita almost 6 times in only 25 years.

from the expression for s^2_{LR}, using a large q appears to increase the efficiency of the procedure reducing the length of the prediction interval.

But large values for q mean that variability in the data associated with higher frequencies is allowed to enter the estimation of low-frequency, long-run variability, thus compromising the major purpose of the procedure. More formally, MW16 argue that the approximation provided by the central limit theorem becomes poor with large values for q and they label this problem as one of robustness (to trade-off with efficiency). Their recommendation is that $q = 12$ is a sensible choice to extract information at low frequencies that is useful to estimate the uncertainty of long-run forecasts.

For a sample of annual data as this one, this means that only cycles with periodicities longer than $\frac{2\pi}{12/\pi T} = \frac{T}{6} = \frac{69}{6} = 11.5$ years are being considered for the long-run forecasts, which appears to be a reasonable selection.

The confidence level. This obviously amounts to the size of corresponding statistical test. Besides the usual $\alpha = 0.05$, some literature considers "one-standard-deviation" intervals, corresponding to a 67% confidence level. This is however inappropriate for the current purposes, where a seemingly simple statistical decision about a "yes or no" question seeks an answer: is it *very unlikely* that country A will converge to country B in the next 82 years? A level $\alpha = 0.33$ would imply a very weak protection to the null. Instead of the "very unlikely" one would have to consider much vaguer, much less precise statements. Hence, as usual, I will report (upper limits for) intervals with $\alpha = 0.01$ and $\alpha = 0.10$ as well.

5 Empirical results

I begin by presenting a preliminary discussion about the meaning of the outcomes of the tests. This is brief but hopefully enlightening. Next I address the empirical results. First for a rather demanding horizon of 42 years, ending in 2060, and later for a much more lengthier time interval of 82 years, ending in 2100. For many of the countries in the sample, this is also more realistic. In the last subsection the horizon will be further extended to 100 years.

5.1 A preliminary clarification

The change in methodology implies also a change in the outcomes of the procedure and hence a change in the interpretation of the results that must be acknowledged from the outset.

The vast majority of the literature on testing for convergence relies on unit root tests and the most common of these takes the unit root, non-stationary behaviour of the gap series, as their null hypothesis. Since non-stationary, unit root or simply $I(1)$ behaviour corresponds to non-convergence and since the classical statistical approach protects the null hypothesis, this means that this method is rather skeptical of convergence. A great deal of evidence is needed to be able to reject the unit root, finding supporting evidence for convergence. The proposed method entails a reversal in the roles of the hypotheses and hence a stance reversal as well. Now it is the convergence hypothesis that is protected, and the lower the size of the test (α) — i.e., the larger the level of confidence $(1-\alpha)$ — the more it is protected. A great amount of evidence is now required to "convict" convergence, leaning towards non-convergence.

This means that the proposed method is much more favourable to convergence than the vast majority of approaches previously adopted in the literature. Therefore, perhaps even more carefully than is usual, non-rejections of the null (of convergence) hypothesis must be clearly understood. Such an outcome, a non-rejection, does not mean that the country's income per capita will certainly converge to the one of the U.S. in h years. It does not even mean that such an event is very likely to occur. It simply means that, on the basis of available information, very strong evidence that the country will not converge cannot be found.

A simple example may help to clarify this issue. Between 1950 and 2018 Argentina has widened (not closed) its gap with the U.S. at an average rate of 0.65% per year. Is this sufficient to consider that its convergence will not occur until 2060 ($h = 42$)? No, with the usual 5% sized test, it is not sufficient. The point forecast is this same value but it is not sufficient to reject convergence because it has an associated large amount of estimated (sampling) uncertainty and also because the usual 5% sized test is built precisely in such a way as to reject the null hypothesis — here, convergence —, when it is true in only 5% of the cases. To reject convergence, the required average rate must be larger than the upper bound of the 95% prediction interval. Convergence is rejected only because this condition is satisfied (see table 1), not because the point forecast provides such a pessimistic outlook that it is even negative. Notice further that if the size of the test were only 1% (confidence level of 99%), the upper bound would increase to 3.09% and, since this is larger than the required average rate of 2.60%, rejection of convergence would cease to be possible.

In summary, this forward looking approach makes finding evidence against convergence much harder than usual, and hence it should not be surprising to find much less of that evidence than in the previous literature. It is optimistic about the future even when the past offers little or no support to be so.

5.2 Forecasts for a 42-year horizon

Table 1 contains the results for the 86 countries in the sample whose income per capita in 2018 was below the one for the U.S. .

The general picture that emerges is a mixed one, much better than the one that is usually transmitted in previous (time series) empirical literature, but still more negative than one would expect on the basis of the radical change in both perspective and decision methodology. Indeed, in a time horizon of 42 years, only 54 of those countries are expected to lie in a very wide "basis of attraction" around the position of the technological leader. But even augmenting this set of countries with those with even higher income than the U.S. (presented in table 2), only 53,1% of the population considered here is represented. Such a picture, with only a little more than half the population, for such a long-time horizon and with such a non-demanding decision rule, can hardly be considered encouraging.

This means that almost half of the world population — most likely possibly

Table 1 — Convergence analysis for $h = 42$ and a confidence level of 95%

country	$\overline{\Delta g_h^*}$	$\overline{\Delta g_U}$	ND?	country	$\overline{\Delta g_h^*}$	$\overline{\Delta g_U}$	ND?
Albania	3.82	3.52	n	Lebanon	3.53	5.40	y
Argentina	2.60	1.84	n	N. Macedonia	3.44	3.70	y
Australia	0.25	1.04	y	Malasya	1.91	3.69	y
Austria	0.60	2.50	y	Malta	1.30	4.68	y
Bangladesh	6.20	2.91	n	Mauritius	2.41	2.93	y
Barbados	3.64	1.95	n	Mexico	2.88	1.85	n
Belgium	0.79	1.14	y	Mongolia	3.38	5.78	y
Bosnia-Herze.	3.97	5.33	y	Montenegro	2.48	6.63	y
Botswana	2.98	5.97	y	Morocco	4.47	1.52	n
Brazil	3.27	3.22	n	Myanmar	5.35	4.29	n
Bulgaria	2.62	3.58	y	Namibia	4.31	1.28	n
Canada	0.50	0.70	y	Netherlands	0.36	1.30	y
Cape Verde	4.98	3.30	n	New Zealand	1.07	0.82	n
Chile	2.18	2.11	n	Nicaragua	5.75	2.29	n
China	3.43	4.76	y	Pakistan	5.49	1.95	n
Colombia	3.35	1.61	n	S. of Palestine	5.63	2.55	n
Congo	5.41	2.47	n	Panama	2.13	3.56	y
C. Rica	3.16	1.81	n	Paraguay	4.24	2.51	n
Croatia	2.19	4.58	y	Peru	3.58	2.37	n
Cuba	4.51	3.23	n	Philippines	4.56	2.43	n
f. Czechos.	1.49	2.60	y	Poland	1.67	3.32	y
Cyprus	1.69	3.46	y	Portugal	1.71	2.48	y
Denmark	0.42	1.13	y	P. Rico	1.09	2.96	y
Dominican R.	2.97	3.68	y	Romania	2.41	5.85	y
El Salvador	4.43	2.00	n	Russian Fed.	1.92	4.05	y
Finland	0.84	1.93	y	Serbia	3.25	5.54	y
France	0.86	1.31	y	Seychelles	1.50	3.71	y
Germany	0.43	2.66	y	Slovenia	1.52	4.02	y
Ghana	6.10	1.87	n	S. Africa	3.61	1.48	n
Greece	2.04	3.28	y	S. Korea	0.90	5.28	y
Guatemala	4.79	1.09	n	Spain	1.34	2.73	y
Honduras	5.70	1.08	n	Sri Lanka	3.71	2.70	n
H. Kong	0.20	3.73	y	Swaziland	4.58	4.62	y
Hungary	1.83	2.71	y	Sweden	0.46	1.10	y
Iceland	0.58	1.62	y	Taiwan	0.51	4.48	y
India	4.99	2.80	n	Thailand	2.86	3.73	y
Indonesia	3.67	3.87	y	Tunisia	3.77	2.10	n
Israel	1.23	2.77	y	Turkey	2.51	2.75	y
Italy	1.13	2.10	y	U.K.	0.89	0.50	n
Jamaica	4.83	2.57	n	f. USSR	2.48	3.50	y
Japan	0.85	4.10	y	Uruguay	2.40	1.95	n
Jordan	3.74	3.59	n	Vietnam	4.99	3.34	n
Lao PDR	5.12	3.65	n	f. Yugosla.	2.87	4.47	y

Notes: h denotes the forecast horizon in years; $\overline{\Delta g_h^*}$ denotes the required average rate of annual decrease of the output gap to the U.S. in a 42-year time horizon and $\overline{\Delta g_U}$ is the estimated maximum attainable reduction rate for that same horizon using a 95% confidence level (for that same country); "ND?" questions whether the country is not forecast to diverge, i.e., whether $\overline{\Delta g_h^*} \leq \overline{\Delta g_U}$.

more than half due their faster growing populations — is forecast to still live in laggard countries in 2060, with living standards much worse than that of U.S. citizens. Furthermore, this group includes such populous countries as Argentina (44.7 million), Bangladesh (167.5) Brazil (211.31) Colombia (48.1), Ghana (27.9), India (1298.1), Morocco(35.5), Myanmar (54.4), Pakistan (220.0), Peru (31.4), Philippines (112.1), Vietnam (97.1) and South Africa (55.3).

Almost all the countries in this falling behind group are from Africa, Asia or Central and South America. However, a rather surprising result is the inclusion of the previous leader, the U.K. (66.7), in this group. This is explained by a generally divergent sample behaviour — the gap augmenting, not decreasing, 0.07% yearly on average — implying also a need for a relatively high required rate, almost as high as the one of South Korea, and by a reasonably small estimate for the measure of forecast uncertainty. A somewhat similar case is the one of New Zealand, whose sample path is even worse, with a gap widening almost 0.5% on average per year. Further still, adding more countries to the sample, excepting both oil exporters as well as (mostly) Eastern European with smaller sample sizes, would most likely worsen, not improve, the general picture.

Still with a somewhat gloomy outlook, several high income countries appear to be affected by a growth fatigue syndrome, with relatively small, albeit sufficient, maximum attainable rates of gap reduction to the U.S.. This is the case of Australia, Belgium, Canada, Denmark, France, the Netherlands and Sweden.

On the other hand, a relatively large group of countries appears capable of closing their gap at very high speed, lending support to the hypothesis. This is the case of Bosnia-Herzegovina, Botswana, China, Croatia, Lebanon, Malta, Mongolia, Montenegro, Romania, Serbia and S. Korea, which are expected to close their gaps at rates that may exceed 4.5% per year; most are Eastern European or Asian.

In table 2 the role of the U.S. is inverted because it is now only one of the followers, behind the new (oil) “leader”, now (pragmatically) Norway by convention since it has the highest income per capita (in PPP units, in 2018), more than 50% higher than the one of the U. S.. Curiously, apparently only Ireland and Singapore

Table 2 — Convergence analysis for very high income countries, with $h = 42$ and a confidence level of 95% (Norway is the new leader)

country	$\overline{\Delta g_h^*}$	$\overline{\Delta g_U}$	ND?	country	$\overline{\Delta g_h^*}$	$\overline{\Delta g_U}$	ND?
Ireland	0.64	1.93	y	Switzerland	0.76	0.46	y
Luxemburg	0.92	0.64	n	U.S.A.	1.01	−0.14	y
Singapore	0.51	4.04	y				

Notes: h denotes de forecast horizon in years; $\overline{\Delta g_h^*}$ denotes the required average rate of annual decrease of the output gap to Norway in a 42-year time horizon and $\overline{\Delta g_h^*}$ is the estimated maximum attainable reduction rate for that same horizon using a 95% confidence level (for that same country); "ND?" questions whether the country is not forecast to diverge, i.e., whether $\overline{\Delta g_h^*} \leq \overline{\Delta g_U}$.

now seem to follow the pace of Norway, Luxemburg and Switzerland clearly lagging behind and the U.S. even forecast to diverge. This behaviour is clearly supported in the past, with the U.S. loosing −1.46% on average per year during the sample period. Given the importance of oil in Norway's economy there appears to exist no reason to worry.

The comparison with previous empirical results based on the unit root (UR) approach confirms the expectations based on the argumentation made at the beginning of this section. Considering the common set of 24 countries previously mentioned, a large difference in outcomes emerges because UR testing allowed only 10 countries to pass the convergence test, while the proposed method is much less demanding, with only New Zealand and the U.K. failing. A further example is that, for instance, Hungary, Israel and Portugal are considered as non-converging countries (in the past) with the UR approach, while they are now viewed as not very unlikely to catch the leader (in the future).

5.3 Forecasts for a 82-year horizon

Table 3 contains the results for the more ambitious exercise with $h = 82$, pushing the forecast horizon to 2100. The level of confidence is maintained at 95%, as usual (i.e., the size of the tests is maintained at 5%). Two opposite forces drive the change in results:

a) a decline in the required average rate of decrease of the income gap, the longer time interval permitting a much slower pace of catching-up;

b) a decline in the estimated maximum attainable rate due to a "law of large numbers effect" (MW16).

As can be easily observed in table 3, the first effect is much stronger than the second (an example is provided in the appendix). Therefore, 15 more countries now pass the test with this much longer time horizon. And among them we now find such populous countries as Argentina, Brazil, Myanmar and Vietnam. However, forecasts still maintain such important countries as Bangladesh, Colombia, Ghana, India, Morocco, Pakistan and Philippines very far from the frontier. And, still very surprisingly, the same is true for the previous leader, the U.K., still predicted to lie much behind the current leader in 2100. At least for about 2100 million people, corresponding to a "n" in the column of the answer to the question in table 3, the outlook for 2100 is still gloomy.

At this point it is perhaps useful to recall that the question is not whether each country is predicted to complete its catch-up process until the end of the forecast horizon. Rather, the question is whether the country is *not forecast to diverge*. Hence, a positive answer means that the country is likely to attain a convergence basin or orbit. It is the uncertainty that is associated with the forecast that allows making such a non-negative forecast and not anything further than that. In other words, attaining that orbit seems likely but there is no guarantee that a "landing" will really follow (nor even that the orbit is really attained).

To reinforce this point, notice that if the confidence level is reduced to 90%, i.e., the size of the corresponding test is increased to 10%, eight countries leave the positive list and reenter again the negative one: Argentina, Cape Verde, Cuba, Mexico, New Zealand, Peru, Uruguay and Vietnam (see table A.1 of the appendix).

Table 3 — Convergence analysis for $h = 82$ and a confidence level of 95%

country	$\overline{\Delta g_h^*}$	$\overline{\Delta g_U}$	ND?	country	$\overline{\Delta g_h^*}$	$\overline{\Delta g_U}$	ND?
Albania	1.96	3.10	y	Lebanon	1.81	4.49	y
Argentina	1.33	1.43	y	N. Macedonia	1.76	3.24	y
Australia	0.13	0.90	y	Malasya	0.98	3.33	y
Austria	0.31	2.26	y	Malta	0.67	4.33	y
Bangladesh	3.17	2.50	n	Mauritius	1.23	2.53	y
Barbados	1.86	1.62	n	Mexico	1.48	1.61	y
Belgium	0.40	1.01	y	Mongolia	1.73	5.24	y
Bosnia-Herze.	2.03	4.59	y	Montenegro	1.27	5.84	y
Botswana	1.53	5.52	y	Morocco	2.29	1.27	n
Brazil	1.67	2.83	y	Myanmar	2.74	3.81	y
Bulgaria	1.34	3.15	y	Namibia	2.21	0.96	n
Canada	0.26	0.60	y	Netherlands	0.19	1.16	y
Cape Verde	2.55	2.91	y	New Zealand	0.55	0.61	y
Chile	1.12	1.77	y	Nicaragua	2.94	1.76	n
China	1.76	4.34	y	Pakistan	2.81	1.73	n
Colombia	1.71	1.36	n	S. of Palestine	2.88	2.11	n
Congo	2.77	2.05	n	Panama	1.09	3.18	y
C. Rica	1.62	1.58	n	Paraguay	2.17	2.11	n
Croatia	1.12	4.05	y	Peru	1.83	1.96	y
Cuba	2.31	2.67	y	Phillipines	2.34	2.10	n
f. Czechos.	0.76	2.26	y	Poland	0.85	2.93	y
Cyprus	0.87	3.13	y	Portugal	0.87	2.27	y
Denmark	0.22	0.98	y	P. Rico	0.56	2.73	y
Dominican R.	1.52	3.31	y	Romania	1.23	5.35	y
El Salvador	2.27	1.67	n	Russian Fed.	0.99	3.53	y
Finland	0.43	1.72	y	Serbia	1.67	4.82	y
France	0.44	1.15	y	Seychelles	0.77	3.38	y
Germany	0.22	2.40	y	Slovenia	0.78	3.57	y
Ghana	3.13	1.46	n	S. Africa	1.85	1.19	n
Greece	1.05	2.93	y	S. Korea	0.46	4.98	y
Guatemala	2.45	0.79	n	Spain	0.69	2.51	y
Honduras	2.92	0.81	n	Sri Lanka	1.90	2.37	y
H. Kong	0.10	3.41	y	Swaziland	2.35	4.18	y
Hungary	0.94	2.41	y	Sweden	0.24	0.96	y
Iceland	0.30	1.44	y	Taiwan	0.26	4.26	y
India	2.56	2.49	n	Thailand	1.46	3.42	y
Indonesia	1.88	3.46	y	Tunisia	1.93	1.89	n
Israel	0.63	2.48	y	Turkey	1.29	2.52	y
Italy	0.58	1.89	y	U.K.	0.46	0.40	n
Jamaica	2.47	2.13	n	f. USSR	1.27	3.09	y
Japan	0.44	3.73	y	Uruguay	1.23	1.58	y
Jordan	1.92	3.04	y	Vietnam	2.55	2.94	y
Lao PDR	2.62	3.28	y	f. Yugosla.	1.47	3.91	y

Notes: h denotes the forecast horizon in years; $\overline{\Delta g_h^*}$ denotes the required average rate of annual decrease of the output gap to the U.S. in a 42-year time horizon and $\overline{\Delta g_U}$ is the estimated maximum attainable reduction rate for that same horizon using a 95% confidence level (for that same country); "ND?" questions whether the country is not forecast to diverge, i.e., whether $\overline{\Delta g_h^*} \leq \overline{\Delta g_U}$.

5.4 Robustness

In this subsection I will shake some pillars of the previous analysis to gauge whether its main conclusions still stand.

Extending the forecast horizon. In table 4 the forecast horizon is extended to 100 years. The results are presented only for those countries still outside the 95% prediction interval with $h = 82$. Needless to say that, although the estimates do not reflect the increased uncertainty with more distant dates, it must be acknowledged that as they are pushed further into the future the likelihood that large shocks occur that derail a convergent trajectory or that pull a divergent path to a convergent one increase as well. Therefore, the results of this exercise must be viewed with much caution.

Table 4 contains a few positive results as the extension of the horizon allows 7 more countries to pass the test. And three of them are particularly relevant due to their population: Philippines, the U.K. and, above all, India. Such results, however, do not appear particularly encouraging, not only due to the change in the nature of the decision method but also to a really very long time distance. In a rapidly changing world, if the convergence hypothesis is meant to really imply practically relevant effects, an 100-year horizon appears too much far off in the future. And even with such a deep horizon and in the best case scenario, countries as Bangladesh, Colombia, Ghana, Morocco and Pakistan are predicted to lie much behind the economic frontier.

Changing the lowest frequency. This amounts to change q and I have experimented with the two values adjacent to the base, 11 and 13, holding the forecast horizon and the level of confidence fixed at 82 years and 95%, respectively. In both cases the changes in the estimated upper bounds are small and produce only a few qualitative changes in outcomes.

The most frequent, though not exclusive effect of reducing q to 11 is to obtain slightly larger estimates for the upper bound of the intervals. Only in the cases of India and Tunisia this is sufficient to anticipate in a few years the arrival to the

Table 4 — Convergence analysis for $h = 100$ and a confidence level of 95%, only for countries outside the 95% prediction interval after 82 years

country	$\overline{\Delta g_h^*}$	$\overline{\Delta g_U}$	ND?	country	$\overline{\Delta g_h^*}$	$\overline{\Delta g_U}$	ND?
Bangladesh	2.60	2.41	n	Jamaica	2.03	2.04	y
Barbados	1.53	1.56	y	Morocco	1.88	1.22	n
Colombia	1.41	1.31	n	Namibia	1.81	0.89	n
Congo	2.27	1.96	n	Nicaragua	2.41	1.65	n
C. Rica	1.33	1.53	y	Pakistan	2.31	1.68	n
El Salvador	1.86	1.60	n	S. of Palestine	2.36	2.02	n
Ghana	2.56	1.37	n	Paraguay	1.78	2.03	y
Guatemala	2.01	0.73	n	Philippines	1.92	2.02	y
Honduras	2.40	0.76	n	Tunisia	1.58	1.85	y
India	2.10	2.43	y	U.K.	0.37	0.39	y

Notes: h denotes the forecast horizon in years; $\overline{\Delta g_h^*}$ denotes the required average rate of annual decrease of the output gap to the U.S. in a 42-year time horizon and $\overline{\Delta g_U}$ is the estimated maximum attainable reduction rate for that same horizon using a 95% confidence level (for that same country); "ND?" questions whether the country is not forecast to diverge, i.e., whether $\overline{\Delta g_h^*} \leq \overline{\Delta g_U}$.

convergence orbit.

When q is increased to 13 the most frequent but not the only effect is to decrease slightly the upper bound of the interval and the only qualitative change concerns, curiously, Tunisia again, whose interval becomes wider, allowing it to anticipate the positive result in a few years.

Changing the level of confidence. That is, automatically and implicitly changing the size of the corresponding test[11]. For this purpose I have considered 90% and 99% forecast intervals. But while reducing this level to only 67%, as is sometimes done, appears inappropriate, augmenting it to 99% does not seem reasonable as well, amounting to an excessively protective position of the convergence hypothesis. On the other hand, reducing it to 90% is not so uncommon and appears adequate in this context. Moreover, it allows improving the power of the tests.

Table A.1 in the Appendix contains both results and again the main message

[11]Recall that a confidence interval with level of $1 - \alpha$ amounts to a test with size α.

still holds, although with some variations. The narrower interval emphasizes several negative features already present in the basic analysis, although somewhat blurred. As previously mentioned, eight more countries do not improve their outlook in relation to the shorter forecast horizon. That is, even in 2100 those countries are still forecast as very likely being considered as laggards. The larger confidence sets obviously provide a much more optimistic view, unfortunately much more unlikely as well.

6 Conclusions

In this paper I propose a new simple method to assess the international income convergence hypothesis. This method retrieves the long-run forecasting definition of convergence by Bernard and Durlauf (1996) and relies on the work of Müller and Watson (2008, 2016, 2020), retaining exclusively the low-frequency variability of the data. Short and medium-term information contents that may not only obscure but even distort the analysis are therefore bypassed. Robustness to non-stationarity issues, including those originating in breaks in the behaviour of the series, is achieved through the application of the method not to the original series but to their first differences.

Application to a sample of 90 different countries corresponding to 5.828 billion people produced mixed evidence, that is more positive to the hypothesis than most empirical studies using time series to assess it and, in particular, more positive than the general idea echoed by Johnson and Papageorgiou (2020) in their influential paper.

Indeed, in a time horizon of 42 years, in 2060, a majority of countries are predicted *not* to lie much behind the technological leader. Considering that they represent more than half of the total sample population in 2018, this result is much more optimistic than in most previous studies (as, e.g., in Lopes, 2023). Moreover, many previous studies finding supporting evidence can be dismissed on the grounds of deficiencies and limitations. For instance, some have considered rather restricted sets of countries, composed of high and medium-high income countries only. Others have considered convergence to a group average, rather than to a technological leader. Still others have considered inadequate alternative hypothesis to the unit root, non-convergence null hypothesis, considering non-stationarity around (unreasonable) deterministic trends, and thereby employing tests with reduced power.

Taking all these weaknesses into consideration, and besides Startz (2020), this is possibly the empirical study which is most favourable to convergence. And yet, many of the results can be hardly considered as encouraging. For instance, almost half of the sample population is forecast to still live in laggard countries in 2060. When the horizon is further extended to 2100, fifteen more countries, some with large populations, are expected to pass the test. But, for more than 2100 million people the outlook for 2100 (a coincidence) is still gloomy. And very surprisingly the previous leader, the U.K., is still predicted to be one of the countries that fail the test. If the convergence hypothesis purports to imply really significant effects on people' standards of living, horizons that lie too much deep in the future resemble mere mirages, the more so the more rapidly the world changes.

In other words, though mixed and more positive than most previous evidence, the one that is gathered here casts many doubts on the inevitability of income convergence, at least in practically relevant time frames and as worldwide phenomenon. Moreover, since convergence was now given the status of the protected hypothesis, becoming much harder to reject, at least some of the dissipated bleakness is due to a reversal of the roles of the hypotheses.

Geographically, the evidence does not present great surprises: a) the region that appears to be more at odds with convergence is Central America; b) most countries in Africa appear to be also facing the bleakest perspectives but there are important exceptions as Botswana and Swaziland; c) in Asia, several very populous countries lagging much behind (Bangladesh, India, Pakistan and Philippines) coexist with some of the fastest growing economies in the world (China, Mongolia, S. Korea and Taiwan); d) finally, Eastern Europe countries appear to be closing their gaps at a very fast pace. On the other hand, I consider the U.K. to be the most important negative surprise: in the best case scenario, not even 82 years appear to be sufficient for it to catch-up with the current leader. The main reason lies not so much on the size of the current gap, which is only slightly larger than Japan's, but rather on its (slowly) diverging sample path.

References

[1] Bernard, Andrew, B., and Steven N. Durlauf (1996). Interpreting tests of the convergence hypothesis, *Journal of Econometrics*, 71, 161-73.

[2] Bolt, Jutta, and Jan L. van Zanden (2020). Maddison style estimates of the evolution of the world economy. A new 2020 update. *Maddison-Project Working Paper* WP-15.

[3] Christopoulos, Dimitris K., and Miguel A. Leon-Ledesma (2011). International output convergence,breaks and asymmetric adjustment, *Studies in Nonlinear Dynamics and Econometrics*, 15 (3), article 4.

[4] Hamilton, James D. (1994), *Time Series Analysis*, Princeton University Press, Princeton.

[5] Hansen, Bruce E. (2022). *Econometrics*, Princeton University Press, Princeton.

[6] Johnson, Paul, and Chris Papageorgiou (2020). What remains of cross-country convergence? *The Journal of Economic Literature*, 58, 129-175.

[7] King, Alan and Carlyn Ramlogan-Dobson (2014). Are income differences within the OECD diminishing? Evidence from Fourier unit root tests, *Studies in Nonlinear Dynamics and Econometrics*, 18 (2), 185-199.

[8] Lopes, Artur S. (2016). A simple proposal to improve the power of income convergence tests, *Economics Letters*, 138, 92 - 95.

[9] Lopes, Artur S. (2023). Non-convergent incomes with a new DF-Fourier test:

most likely you go your way (and I'll go mine), working paper, ISEG-ULisboa.

[10] Milanovic, Branko (2016). *Global Inequality: a new approach for the age of globalization*, The Belknap Press of Harvard University Press.

[11] Müller, Ulrich K., James H. Stock and Mark W. Watson (2020). An econometric model of international long-run growth dynamics, NBER working paper 26593.

[12] Müller, Ulrich K., and Mark W. Watson (2008). Testing models of low-frequency variability, *Econometrica*, 76, 5, 979 - 1016.

[13] Müller, Ulrich K., and Mark W. Watson (2016). Measuring uncertainty about long-run predictions, *Review of Economic Studies* 83, 1711 - 40.

[14] Müller, Ulrich, and Mark W. Watson (2020). Low-Frequency Analysis of Economic Time Series, draft chapter for *Handbook of Econometrics*, Volume 7, edited by S. Durlauf, L.P. Hansen, J.J. Heckman, and R. Matzkin.

[15] Pesaran, Mohammad H. (2007). A pair-wise approach to testing for output and growth convergence, *Journal of Econometrics*, 138, 312-55.

[16] Quah, Danny (1993). Galton's Fallacy and Tests of the Convergence Hypothesis, *Scandinavian Journal of Economics*, 95(4), 427-443.

[17] Startz, Richard (2020). The next hundred years of growth and convergence, *Journal of Applied Econometrics*, 35, 99-113.

[18] The Economist (2014). Economic convergence: the headwinds return, Sep 13th (Updated Jun 11th 2015).

7 Appendix: brief methodological review

This appendix presents with some more detail the method proposed in Müller and Watson (2016, MW16) to measure uncertainty in (very) long-term forecasts and to build confidence intervals for the forecasts using these measures. It is one of the simplest problems addressed in MW16, not only because it is a case of univariate analysis but also because it concerns a covariance stationary, $I(0)$, time series. I provide also a few brief examples of the forecast interval for the particular case at hand.

The low-frequency methods of analysis for time series are an alternative to low-frequency band pass (or simply lowpass) filters, which consist in projecting the data on a small number of low frequency trigonometric series. A very well known and frequently used lowpass filter is the so-called "Hodrick-Prescott" filter and the goal here is similar: to isolate or to extract from the data only that component that corresponds to low-frequency, long term variation, willfully neglecting medium and high frequency variation, as these may be harmful to the purposes of the study. MW16 label this property of low-frequency methods simply as robustness.

This is achieved estimating with OLS the regression of the series, x_t, on some low-frequency deterministic functions of time, the trigonometric terms $\cos[(t - 1/2)\pi j/T]$ for $j = 0, 1, \ldots, q$. These are associated with the frequencies $0, \pi/T, 2\pi/T, \ldots, q\pi/T$. Recalling the inverse relation between frequency, ω (in radians), and period, given by

$$\text{period} = 2\pi/\omega,$$

then, besides the sample mean, which is associated with the zero frequency, the idea consists in capturing the variation associated with cycles with periodicity $2T, T, 2T/3, \ldots, 2T/q$. For instance, with annual data, $T = 60$ and $q = 12$, the variations with the highest frequency, i.e., with shortest periodicity, are those corresponding to cycles which are 10 $(= 120/12)$ years long, which is clearly beyond the usual periodicities associated with business cycles (between 1.5 and 8 years). The first coefficient estimate of this regression is simply the sample mean ($\bar{x}_{1:T} = T^{-1} \sum_{t=1}^{T} x_t$) and the other coefficient estimates are labeled the (discrete) cosine transforms of the series and represented with $\mathbf{X}_{T,1:q}$ in MW16 or, preferably, with $\mathbf{X}_T$ in Müller and Watson (2020). Instead of this notation, in the main text I stick with the traditional OLS representation, $\widehat{\beta}^q_{LF}$. All the coefficient estimates are collected in the vector $\mathbf{X}_T^0 = (\bar{x}_{1:T}, \mathbf{X}_T')'$.

The goal here is to forecast the value of $\bar{x}_{T+1:T+h}$, which represents the average value for x_t from time period $T+1$ through time $T+h$, h representing the forecast horizon. For this purpose, a predictive distribution for $\bar{x}_{T+1:T+h}$ is needed to build a confidence interval. MW16 prove that the preditive distribution is

$$\frac{(\bar{x}_{T+1:T+h} - \bar{x}_{1:T})}{\sqrt{(h^{-1} + T^{-1})s_{LR}^2}}|\mathbf{X}_T^0 \overset{a}{\sim} \text{Student} - t_{(q)},$$

where $s_{LR}^2 = T\mathbf{X}_T'\mathbf{X}_T/q$ is an estimator for the long-run variance (denoted with $s_{LR}^2 = (T/q)\widehat{\beta}_{LF}^{q\prime}\widehat{\beta}_{LF}^q$ in the main text). Its particularly simple form is due to the orthogonality of the regressors.

Therefore, to obtain a two-sided $100(1-\alpha)\%$ confidence interval for the forecast one has to get $t_{(1-\alpha/2)}^q$, the $(1-\alpha/2)$ quantile from the Student-t distribution with q degrees of freedom, allowing to write

$$\Pr\left[-t_{(1-\alpha/2)}^q < \frac{\bar{x}_{T+1:T+h} - \bar{x}_{1:T})}{\sqrt{(h^{-1} + T^{-1})s_{LR}^2}} < t_{(1-\alpha/2)}^q\right] = 1 - \alpha.$$

Solving the inequality for $\bar{x}_{T+1:T+h}$ results in the interval with endpoints

$$\bar{x}_{1:T} \pm t_{(1-\alpha/2)}^q \left(1 + \frac{T}{h}\right)^{1/2} T^{-1/2} s_{LR},$$

which is reproduced in (2) with $\overline{\Delta g_T}$ in the place of $\bar{x}_{1:T}$ since I work with the Δg_{kt} series. For purposes of statistical testing, in case the alternative hypothesis were the usual two-sided hypothesis, the rejection or critical region of the test would be the complementary set of this interval.

But in the present problem there is no evidence against convergence when country k exceeds the upper limit of the interval, that is, when it reduces its gap with the leader at a rhythm that is faster than what is needed to close the gap. This is a case where the data does not contradict the convergence hypothesis. Therefore, a one-sided interval with its corresponding one-sided rejection region (with only the unfavourable outcomes for convergence) must be employed instead. In other words, the evidence goes against the convergence hypothesis when the required average rate exceeds the upper limit of the one-sided confidence interval. This interval, or non-rejection region, is given in (3), for a size (α) of the test. Its complementary set is the rejection or critical region of the test.

As a first example, consider the case of Argentina in table 1: a forecast for 2160, i.e., $h = 42$, and a test with size $\alpha = 0.05\%$. The gap in 2018 was 36.779(= 55.335 − 18.556) in PPP units, i.e., −1.09259 in logged terms. To make this gap disappear in 42 years it must decrease at an average rate of 2.06% ($= \overline{\Delta g_h^*}$) per year. Instead, the point forecast is that the gap continues to increase because it was what happened during the sample period: it has increased at an average rate of about 0.65% on average per year. However, due to the uncertainty associated with this forecast, and with a 95% confidence level, the upper limit of the one-sided forecast interval is 1.84% ($= \overline{\Delta g_U}$) . In case it is realized, it would represent a rather positive evolution, at least much better than the past. Nevertheless, it would be insufficient to close the gap in 2060. Hence, in this case the evidence is clearly against the convergence hypothesis.

However, if the forecast horizon is extended to 2100 ($h = 82$), the average rate that is needed is only 1.33% per year, that is, only a bit more than half the previous desired rate. Although the point forecast for the average is the same, the increase in h implies a reduction in the estimated uncertainty and the upper limit for the prediction interval is now 1.43%, which is more than sufficient to eliminate the gap. Therefore, still with a 5% size for the test, the evidence is not unfavorable for convergence when the horizon is extended to 2100.

Consider now the case of Botswana, which is marked by a remarkable evolution during the sample period: from only 3.2% in 1950, Botswana managed to arrive at 2018 with 28.6% of U.S. per capita income. Considering an horizon of 42 years, the average rate that is needed is ($\overline{\Delta g_h^*}$ =)2.98% per year. But already in the sample period Botswana managed to reduce its gap with the U.S. at the tremendous rate of 3.22% per year, and this is the point forecast. This would be already sufficient but the upper limit for the 95% confidence interval, which incorporates the estimated uncertainty, is even much higher, at 5.95% (= $\overline{\Delta g_U}$). There is thus no evidence against convergence for this horizon and, as can be easily observed for $h = 82$ (table 2), although the the threshold of the interval is lower, at 5.52%, the required average rate is much lower, at only 1.53%, and hence, still with a 5% size test, the supporting evidence for the convergence hypothesis is now even more comfortable.

8 Appendix: further results with 90% and 99% confidence intervals

Table A.1 — Convergence analysis for $h = 82$ at 90% and 99% levels

		90% confi.		99% confi.				90% confi.		99% confi.	
country	$\overline{\Delta g_h^*}$	$\overline{\Delta g_U}$	ND?	$\overline{\Delta g_U}$	ND?	country	$\overline{\Delta g_h^*}$	$\overline{\Delta g_U}$	ND?	$\overline{\Delta g_U}$	ND?
Albania	1.96	2.59	y	4.18	y	Lebanon	1.81	3.37	y	6.84	y
Argentina	1.33	0.93	n	2.48	y	N. Macedonia	1.76	2.68	y	4.43	y
Australia	0.13	0.74	y	1.24	y	Malasya	0.98	2.89	y	4.26	y
Austria	0.31	1.96	y	2.88	y	Malta	0.67	3.90	y	5.23	y
Bangladesh	3.17	2.00	n	3.56	y	Mauritius	1.23	2.05	y	3.56	y
Barbados	1.86	1.23	n	2.46	y	Mexico	1.48	1.32	n	2.23	y
Belgium	0.40	0.85	y	1.35	y	Mongolia	1.73	4.58	y	6.63	y
Bosnia-Herze.	2.03	3.67	y	6.52	y	Montenegro	1.27	4.87	y	7.89	y
Botswana	1.53	4.97	y	6.68	y	Morocco	2.29	0.97	n	1.91	n
Brazil	1.67	2.34	y	3.85	y	Myanmar	2.74	3.23	y	5.04	y
Bulgaria	1.34	2.63	y	4.25	y	Namibia	2.21	0.56	n	1.79	n
Canada	0.26	0.48	y	0.86	y	Netherlands	0.19	0.99	y	1.52	y
Cape Verde	2.55	2.43	n	3.91	y	New Zealand	0.55	0.35	n	1.16	y
Chile	1.12	1.36	y	2.63	y	Nicaragua	2.94	1.12	n	3.12	y
China	1.76	3.84	y	5.42	y	Pakistan	2.81	1.45	n	2.31	n
Colombia	1.72	1.07	n	1.99	y	S. of Palestine	2.88	1.58	n	3.23	y
Congo	2.77	1.53	n	3.14	y	Panama	1.09	2.73	y	4.14	y
C. Rica	1.62	1.29	n	2.18	y	Paraguay	2.17	1.63	n	3.12	y
Croatia	1.12	3.40	y	5.43	y	Peru	1.83	1.46	n	3.01	y
Cuba	2.31	1.98	n	4.12	y	Philippines	2.34	1.69	n	2.95	y
f. Czechos.	0.76	1.85	y	3.12	y	Poland	0.85	2.45	y	3.92	y
Cyprus	0.87	2.71	y	4.00	y	Portugal	0.87	2.01	y	2.81	y
Denmark	0.22	0.79	y	1.36	y	P. Rico	0.56	2.44	y	3.33	y
Dominican R.	1.52	2.87	y	4.26	y	Romania	1.23	4.74	y	6.63	y
El Salvador	2.27	1.27	n	2.51	y	Russian Fed.	0.99	2.84	y	4.98	y
Finland	0.43	1.47	y	2.25	y	Serbia	1.67	3.93	y	6.70	y
France	0.44	0.96	y	1.55	y	Seychelles	0.77	2.97	y	4.24	y
Germany	0.22	2.08	y	3.08	y	Slovenia	0.78	3.00	y	4.76	y
Ghana	3.13	0.96	n	2.51	n	S. Africa	1.85	0.84	n	1.93	y
Greece	1.05	2.49	y	3.84	y	S. Korea	0.46	4.61	y	5.74	y
Guatemala	2.45	0.43	n	1.56	n	Spain	0.69	2.23	y	3.09	y
Honduras	2.92	0.49	n	1.50	n	Sri Lanka	1.90	1.97	y	3.22	y
H. Kong	0.10	3.03	y	4.23	y	Swaziland	2.35	3.64	y	5.32	y
Hungary	0.94	2.03	y	3.19	y	Sweden	0.24	0.79	y	1.33	y
Iceland	0.30	1.21	y	1.91	y	Taiwan	0.26	3.99	y	4.83	y
India	2.56	2.12	n	3.27	y	Thailand	1.46	3.04	y	4.21	y
Indonesia	1.88	2.96	y	4.51	y	Tunisia	1.93	1.64	n	2.43	y
Israel	0.63	2.14	y	3.22	y	Turkey	1.29	2.25	y	3.10	y
Italy	0.58	1.62	y	2.45	y	U.K.	0.46	0.29	n	0.65	y
Jamaica	2.47	1.60	n	3.25	y	f. USSR	1.27	2.41	y	4.52	y
Japan	0.44	3.28	y	4.69	y	Uruguay	1.23	1.12	n	2.55	y
Jordan	1.92	2.38	y	4.45	y	Vietnam	2.55	2.44	n	3.98	y
Lao PDR	2.62	2.82	y	4.24	y	f. Yugosla.	1.47	3.22	y	5.36	y

Notes: h denotes the forecast horizon in years; $\overline{\Delta g_h^*}$ denotes the required average rate of annual decrease of the output gap to the U.S. in a 82-year time horizon and $\overline{\Delta g_U}$ is the estimated maximum attainable reduction rate for that same horizon with the specified confidence level; "ND?" questions whether the country is not forecast to diverge, i.e., whether $\overline{\Delta g_h^*} \leq \overline{\Delta g_U}$.