Nonlinear Economic Growth: Some Theory and Cross-Country Evidence (with appendices)*

Davide Fiaschi   Andrea Mario Lavezzi
University of Pisa   University of Pisa

March 9, 2006

Abstract

This paper aims to test the existence of different growth regimes, that is of different relationships between growth rate and income level. We propose a simple nonlinear growth model and test its empirical implications by estimating Markov transition matrices and stochastic kernels. We show that growth is indeed nonlinear: a first phase of slow or zero growth is followed by a take-off and, finally, by a phase of deceleration. We discuss the relevance of these results with respect to the issue of convergence and reversibility of development, in the light of models of structural change and technological diffusion.

Keywords: nonlinear growth, distribution dynamics, convergence, structural change, technological diffusion.

JEL classification numbers: O11, O40, C14, C21.

---

*Corresponding author: Davide Fiaschi, Dipartimento di Scienze Economiche, Via Ridolfi 10, 56124 Pisa, Italy, e-mail: dfiaschi@ec.unipi.it, phone/fax: ++39.050.2216208/++39.050.598040. Two anonymous referees provided valuable suggestions on the submitted draft. We also thank Carlo Bianchi, Eugene Cleur, Steven Durlauf, Raffaele Paci and Francesco Pigliaru for helpful comments. Usual disclaimers apply.
1 Introduction

This paper discusses the issue of nonlinearities in economic growth. Different theories suggest that economic growth is a nonlinear process (see Lewis (1956), Rostow (1960), Mas-Colell and Razin (1973), Murphy et al. (1989), Peretto (1999) and Galor and Weil (2000)). According to this approach, the growth path of an economy displays an initial phase of stagnation, followed by a take-off in which growth rates are increasing, and eventually reaches a regime of steady growth. These different growth regimes, associated to different levels of development, are generated by the structural transformations faced by a growing economy. This also implies that an economy not showing the proper conditions for the take-off can remain trapped in a long period of stagnation. In this framework the influence of international technological spillovers on the growth process, especially on the take-off, is generally negligible, with respect to the internal sources of accumulation.

Differently, another important strand of research focuses on distinct kinds of interactions which may take place among economies. This literature devoted particular attention to technological spillovers (see e.g. Parente and Prescott (1994) and Basu and Weil (1998)). These contributions analyse the conditions that allow a country, starting its economic development, to benefit from the knowledge accumulated by richer countries, and therefore increase its growth rate. Lucas (2000) provides a very simple model of this process. In this setting a nonlinear growth path could be the result of different adoption speeds, a feature not present in Lucas (2000), in particular when adoption speed increases as a country develops.

In a previous paper (Fiaschi and Lavezzi (2003)) we studied nonlinearities in growth and convergence adopting the distribution dynamics approach. There we followed the current literature and used relative per capita GDP, that is incomes expressed with respect to world average income. The underlying justification for this normalization is the existence of a world trend of technological progress which benefits all countries.

\footnote{1}{For example Peretto (1999) argues that a nonlinear growth process is the result of the transition from growth generated by capital accumulation, subject to decreasing returns to scale, to growth based on knowledge accumulation.}

\footnote{2}{In particular, Basu and Weil (1998) focus on the concept of “appropriate technology”. They argue that technological progress can be hampered or slowed down because technology is specific to capital/labour ratios (capital includes both physical and human capital), so that the technology of leader countries cannot quickly diffuse in backward countries. Alternatively, Parente and Prescott (1994) suggest that various type of barriers (e.g. legal) may check the international diffusion of technology.}

\footnote{3}{Lucas also mentions a reason why poor countries can grow more than rich countries: the flows of resources from rich to poor countries due to diminishing returns to accumulation.}
In this paper we take a different perspective. Our aim is to detect the possible nonlinear growth dynamics of a country related to its own development process (i.e. there may exist threshold effects not related to other countries’ growth). Therefore the (absolute) level of development of a country is the key variable and in this regard we consider the absolute level of per capita GDP, without any reference to a world technological trend. By this choice of the data, we can address some issues which could not be studied in Fiaschi and Lavezzi (2003).

In particular we refer to: i) the identification of a take-off path, and of its shape, as distinct from the problem of catching-up, the latter being represented in terms of relative incomes by a shrinking of the world income distribution around the value of 1. In other words, the empirical predictions of a model such as Lucas (2000), where the issue is whether, when and in which way a country starts its growth process, cannot be satisfactorily tested with data on relative GDP. ii) The possible presence of poverty traps in absolute terms, that is of absolute levels of deprivation in which a country tends to persist. This would raise different concerns with respect to poverty interpreted as economic distance from leading countries. iii) The possibility to shed light on the reversibility of the growth process, which also takes a different meaning than when interpreted as a reversal of relative positions due to leapfrogging (see e.g. Kremer et al. (2001), who find the growth process as basically irreversible). In addition, in this paper we utilize a different database: we consider 122 countries from 1950 to 1998 from Maddison (2001), while in Fiaschi and Lavezzi (2003) we considered 120 countries from 1960 to 1989 from the Penn World Table 5.6. The longer period allows our data set to span a larger GDP range.

From the methodological point of view we extend the empirical analysis of Fiaschi and Lavezzi (2003) in different respects: i) we present a refined analysis of nonlinearities, based on bootstrap inference in the nonparametric estimation of the growth path; ii) we provide an estimation of the long-run distribution in a continuous GDP state space; iii) we offer a discussion of different methods to identify the GDP thresholds separating growth regimes.

We show that growth appears indeed as a nonlinear process. We identify three growth regimes characterized by a different relation between growth rate and income. At low income levels the relation is negative or flat, at intermediate levels it is positive and, finally, at high income levels it is again negative. In particular, countries in the intermediate income range appear to experience rapid growth (take-off) with increasing

\[^{4}\text{In addition, with absolute data, we would not be able to observe whether the GDP differences tend to reduce or not for countries that have crossed the highest GDP threshold, to be defined below.}\]
growth rates. This contrasts with Barro and Sala-i-Martin (2004)’s claim that no evidence supports the hypothesis of “a middle range of values of \( k \) [capital]...for which the growth rate, \( \dot{k}/k \), is increasing in \( k \) and, hence, in \( y \) [income]” (p. 77). However, our description of long-run growth behaviour fits the facts only for a subset of countries, as another subset appears in a persistent state of poverty. At this stage we cannot discriminate whether this phenomenon is permanent or temporary, even if the number of countries in our sample showing a slow/negative growth in per capita GDP is not negligible.

As in Fiaschi and Lavezzi (2003) we study the distribution dynamics with a novel definition of the state space, jointly taking into account income levels and growth rates, which allows us to capture the presence of nonlinearities. The use of Markov transition matrices to study the shape of the growth process with data on the absolute level of per capita GDP requires particular attention in the interpretation of the results. The definition of income classes with per capita GDP can make the conclusions on the long run questionable if one considers that, in the long run, all countries could benefit from a common technological trend, a possibility that we do not explicitly assume but that cannot be completely ruled out. From an empirical point of view, the main piece of evidence against the presence of a such trend is that many African countries show no growth in our period of observation of 50 years.

As a theoretical framework for the empirical analysis, we present a nonlinear growth model close in spirit to Romer (1986) and Zilibotti (1995). Moreover, our paper is related to other works on nonlinearities in economic growth, such as Durlauf and Johnson (1995), Liu and Stengos (1999) and Kalaitzidakis et al. (2001), although these papers focus on the nonlinear effects of some explanatory variables. In particular, Liu and Stengos (1999) and Kalaitzidakis et al. (2001) provide some evidence consistent with our result on the existence of an income range where the growth rate is increasing, but their analysis regards initial income levels.

The rest of the paper is organized as follows: Section 2 contains the preliminary graphical analysis; Section 3 discusses the empirically testable implications of a growth model with nonlinearities; Section 4 reports the quantitative results; Section 5 concludes.
2 Graphical Analysis

In our empirical analysis we relate (log) per capita GDP to growth rates. Data are from Maddison (2001) and refer to 122 countries for the period 1950-1998. In Figure 1 we plot the annual growth rate against per capita GDP for all observations in the sample, and a nonparametric estimation of the relationship between these two variables. The grid represents the state space of growth rate and GDP classes that we use to study the distribution dynamics as a Markov chain. The problem of the identification of the classes’ limits is discussed in Section 4 and in Appendix B.

The nonparametric regression in Figure 1 identifies a nonlinear relationship between the growth rate and the level of GDP. In particular, growth at low levels of GDP is initially high but decreasing, quickly reaching a minimum. Then the relation

---

5 Appendix A contains the country list.
6 For all the nonparametric estimates we used R (2005), in particular the statistical package mgcv, see Wood (2004). 95% confidence bands in Figure 1 are calculated by an appropriate resampling method (wild bootstrap), suggested by Härdle et al. (2004), p. 127. Data sets and codes used in the empirical analysis are available on the authors’ websites (http://www-dse.ec.unipi.it/fiaschi and http://www-dse.ec.unipi.it/lavezzi).
with GDP becomes positive. The estimated growth path, after reaching a peak, shows a tendency to decrease. The larger confidence bands at low GDP levels depend on the high growth volatility. This makes the first decreasing part of the estimate statistically non significant at 5% level, and therefore we cannot exclude that the path in that GDP range is flat (see also Figure 2 below). On the contrary, the inverse U-shape of the growth path in subsequent GDP classes appears statistically significant although, given the increase in the confidence band at the highest GDP levels, we cannot rule out that the path in this range becomes flat at a positive value, as would be predicted by an endogenous growth model.

In Figure 1 we do not control for cross-country heterogeneity. However, Kalaitzidakis et al. (2001) estimate a semiparametric regression in which they consider as explanatory variables for the GDP growth rate the population growth rate, the investment ratio and initial income, allowing the latter to enter nonlinearly (they use a smaller set of countries for a shorter time period). They find a similar shape for the relationship between growth rate and income (see their Figure 1).

By looking only at Figure 1 one may conclude that every country tends to grow in the long run, as the estimate always lies above the x-axis. However, this conclusion is questionable if we consider the performances of many African countries in the sample, which appear stagnating at low GDP levels. The forty-two African countries in the sample had an average growth rate of 1.0% over the period, against an average 2.2% of non-African countries. Within the set of African countries, eight countries had a negative growth rate and eighteen countries had a growth rate lower than 0.5% (half of the group average).

The simple model by Lucas (2000) is compatible with the evidence in Figure 1 for what concerns the aggregate picture, but not for the behaviour of individual countries (in general, this remark holds for all models which consider relative backwardness as a possible advantage for a country). In particular, Lucas’ model predicts that, when lagging countries start growing, they are expected to jump to a very high growth rate, which eventually converges to the growth rate of the leading countries. For African countries, it seems that the growth process started (overall their income grew by 60% in the period), but their income path is generally very volatile and on average rather flat.

From a methodological point of view, the pooling of cross-country data can mislead the researcher in the identification of actual growth patterns, as it identifies a “repre-

\[\text{The dispersion of observations in Figure 1 suggests an inverse relationship between growth volatility and the level of GDP. We investigate this aspect in Fiaschi and Lavezzi (2005).}\]
sentative” growth path. In this regard we will see that the analysis based on transition matrices and stochastic kernels allows us to avoid such mistakes, since it permits us to keep track of the growth path of each individual country.

3 Theoretical framework

In this section we present a simple growth model that can account for the dynamics identified in the previous graphical analysis and derive its empirical implications, which are tested in Section 4.

We consider a simple Solovian growth model with no exogenous technological progress, in which the production function exhibits increasing returns to scale within a certain income range (Appendix D describes the model in more details). Figure 2 depicts this economy under the assumptions that average capital productivity does not decrease so much to generate a poverty trap, and remains sufficiently high for high levels of capital to ensure positive growth in the long run.8

8This model differs from the nonlinear growth model in Fiaschi and Lavezzi (2003), which features common exogenous technological progress for all countries.

Figure 2: endogenous growth model

In Figure 2, $\dot{y}$ represents the level of per capita income and $\ddot{y}$ its growth rate. The
The theoretical framework posits that there are no equilibria and per capita income tends to grow indefinitely. There exists a region of increasing growth rate with respect to income, i.e. \([\bar{y}_I, \hat{y}]\) and two regions of decreasing growth rates, i.e. \([0, \bar{y}_I]\) and \([\hat{y}, \infty]\). Trajectories A and B indicate alternative growth paths at low income levels. In particular, trajectory A represents a flat growth path, while trajectory B generates a poverty trap in GDP class I. The empirical relevance of these trajectories is discussed in Section 4.

This type of dynamics may be generated by different mechanisms. Traditional development theories (Rosenstein-Rodan (1943), Lewis (1956) and Rostow (1960)) emphasized structural change without a formal analysis, provided more recently by Mas-Colell and Razin (1973) and Murphy et al. (1989). Recent contributions include also Zilibotti (1995), who focuses on externalities in the capital accumulation process, Peretto (1999), who analyses the change in the growth rate of a country when investing in R&D becomes the main source of growth, and Galor and Weil (2000), who model the interactions between demographic transition and human capital accumulation in the transition from stagnation to growth. All of these models are based on internal factors and do not consider international technological spillovers.

There also exists a literature on technological catch-up which emphasizes the advantage of backward countries to benefit from technological leaders. In Lucas (2000)’s model, the shape of the growth process appears flat for low-income countries which do not use the “leading” technology. When a country starts to benefit from international technological spillovers (and this happens for all countries in Lucas (2000)), its growth rate immediately jumps to a level that is initially greater than the one of rich countries (the difference is a function of the income gap), and eventually converges to the growth rate of the leaders (see also Howitt and Mayer (2005)). Empirically, we should observe a monotonic growth path which follows the initial “jump”, and irreversibility of development. In contrast, in our model stagnation is followed by a phase of increasing growth rates. This could be the result of a variable adoption speed, which increases as a country develops. Another difference remains in the behaviour of countries in the initial phase of development: while in the model of technological diffusion development is irreversible, in our model the initial development could lead to long-run stagnation in presence of a poverty trap (see trajectory B in Figure 2).

In Figure 2 we superimposed on the space \(\left(\frac{\dot{y}}{y}, y\right)\) a partition in 12 regions based on three levels of per capita income, \(\bar{y}_I, \bar{y}_{II}\) and \(\bar{y}_{III}\) (which define four income classes), and two levels of growth rates, \(g_+\) and \(g_{++}\) (which define three growth rate classes). We choose the growth rate classes in order to include the long-run growth rate in the central class. Since we aim to find empirically testable implications of this model, and
the number of available observations is limited, we must maintain the number of states as low as possible.

The partition in Figure 2 appears well suited to this purpose. In fact, it identifies regions in the space \( \left( \frac{\dot{y}}{y}, y \right) \) characterized by specific relationships between income level and growth rate. In particular, countries in income class I should show decreasing growth rates, while countries in income class II should show low but increasing growth rates. Countries in income class III should show increasing (at least up to \( \dot{y} \)) and/or persistently higher growth rates, with respect to all the other income classes. Finally, countries in income class IV should show decreasing growth rates, which could tend to settle at a medium level. Moreover, we expect countries in income classes I, II and III to show a tendency to move into income class IV in the long run.

In Figure 2 the trajectory represented by the continuous line indicates the case with no poverty traps. As shown in Appendix D, the model can also include this feature if the growth path cuts the \( x \)-axis from above in income class I and from below in income class II (trajectory \( B \) in the figure). At first glance, the latter case does not appear to be in agreement with the nonparametric regression in Figure 1. However, it is well-known that nonparametric estimations underestimate the troughs (see Härdele et al. (2004), p. 47). At any rate, the partition in Figure 2 allows for testing the plausibility of the three alternative trajectories. In particular, in presence of a poverty trap the dynamics of the cross-country income distribution would show a tendency to polarize in classes I and IV and this is the crucial point for testing the relevance of trajectory \( B \). Otherwise, all countries would show a tendency to converge in class IV.

4 Empirical analysis

In this section we discuss our methodology for the empirical investigation and present the results. Following Quah (1993), the growth dynamics of the sample is first represented by Markov transition matrices. In the literature on distribution dynamics, the

\[ \hat{p}_{ij} = \frac{n_{ij}}{n_i} \]

where \( n_i \) is the number of observations in state \( i \), and \( n_{ij} \) is the number of observed transitions from state \( i \) to state \( j \). These estimates, as shown for instance in Anderson and Goodman (1957), are the maximum likelihood estimators of the true transition probabilities \( p_{ij} \). See Appendix E for more details on the properties of transition probabilities’ estimates.
state space has been defined so far only in terms of income levels. In this paper we follow the approach proposed in Fiaschi and Lavezzi (2003) and define the state space in terms of both income levels and growth rates (see Fiaschi and Lavezzi (2003) for further discussion on merits and drawbacks of this definition).

### 4.1 Definition of the state space

In Section 3 we showed that a definition of 3 growth rate classes and 4 income classes, i.e. a total of 12 states, is sufficient to generate empirically testable implications of the presence of nonlinearities. In particular, we adopt the definition of the state space in Table 1.

<table>
<thead>
<tr>
<th>Log of GDP \ Growth rate</th>
<th>&lt; 0.5%</th>
<th>0.5% – 2.5%</th>
<th>&gt; 2.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 – 6.84</td>
<td>I-</td>
<td>I+</td>
<td>I++</td>
</tr>
<tr>
<td>6.84 – 8.29</td>
<td>II-</td>
<td>II+</td>
<td>II++</td>
</tr>
<tr>
<td>8.29 – 9.21</td>
<td>III-</td>
<td>III+</td>
<td>III++</td>
</tr>
<tr>
<td>&gt; 9.21</td>
<td>IV-</td>
<td>IV+</td>
<td>IV++</td>
</tr>
</tbody>
</table>

Table 1: State space definition

The procedure for defining the classes’ boundaries consists in two steps. First we set the growth rate classes on the basis of an estimate of the long-run growth rate, and then we set the GDP classes by a comparison of Figure 1 with Figure 2.

As regards the growth rate classes, the theoretical model shows that the class limits should be set in order to obtain a central class which includes the long-run growth rate. According to our model, the identification of such rate concerns the subset of the wealthiest countries (see Section 2). From an inspection of Figure 1 it seems that the richest countries’ growth rates are in a range whose average is approximately equal to 1.5%.\(^{11}\) Hence, a range of $\pm 1\%$ should reasonably include the long-run growth rate. The three resulting growth rate classes are consequently defined as $[(−∞, \ 0.5\%), \ (0.5\%, \ 2.5\%), \ (2.5\%, \ +∞\%)]$.\(^{12}\) For simplicity we will indicate the three levels of growth rates as “low”, “medium” and “high”.

---

\(^{10}\)Slightly different definition of classes’ boundaries do not affect our results.

\(^{11}\)The value of 1.5% is representative of the average growth rate of the top tail of the distribution of the GDP data (in particular of the top 13% of observations).

\(^{12}\)In Fiaschi and Lavezzi (2003) the growth rate classes are $[(−∞, \ 0.8\%), \ (0.8\%, \ 2.8\%), \ (2.8\%, \ +∞\%)]$. In that case the medium class was interpreted as containing the common long-run exogenous growth rate. This classification is similar to the one found in Jones (1997).
4.2 Results

The definition of per capita GDP classes directly follows from imposing these growth rate classes on Figure 1, taking into account the representation of the state space in Figure 2. In particular, the first GDP class is defined on the basis of minimum reached by the nonparametric estimate. Possible alternatives to identify growth regimes and then define GDP classes are offered by Breiman et al. (1984) and Hansen (2000). In Appendix B we compare these methods with the use of nonparametric estimations, and explain our choice of the latter in this paper.

In the estimate we consider three-year transitions (i.e. from \((y_t, g_t)\) to \((y_{t+3}, g_{t+3})\)) in order to circumvent the possible problem of autocorrelation of shocks. This is particularly relevant for low-income countries, where measurement error can induce serial correlation between growth rates. In Appendix C we report the results with 1-year transitions, which are typically considered in the literature on distribution dynamics, and with 3-year average growth rates, as another way of avoiding the problem of autocorrelation. Overall, our results do not seem to be affected by this phenomenon.

4.2 Results

Table 2 contains the transition matrix obtained by applying the definition of states in Table 1.

<table>
<thead>
<tr>
<th>N. Obs</th>
<th>States</th>
<th>I-</th>
<th>I+</th>
<th>I++</th>
<th>II-</th>
<th>II+</th>
<th>II++</th>
<th>III-</th>
<th>III+</th>
<th>III++</th>
<th>IV-</th>
<th>IV+</th>
<th>IV++</th>
</tr>
</thead>
<tbody>
<tr>
<td>396</td>
<td>I-</td>
<td>0.46</td>
<td>0.19</td>
<td>0.29</td>
<td>0.03</td>
<td>0.02</td>
<td>0.02</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>333</td>
<td>I+</td>
<td>0.27</td>
<td>0.45</td>
<td>0.22</td>
<td>0.02</td>
<td>0.03</td>
<td>0.02</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>387</td>
<td>I++</td>
<td>0.22</td>
<td>0.17</td>
<td>0.41</td>
<td>0.06</td>
<td>0.04</td>
<td>0.10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>870</td>
<td>II-</td>
<td>0.04</td>
<td>0.01</td>
<td>0.01</td>
<td>0.43</td>
<td>0.24</td>
<td>0.28</td>
<td>0.01</td>
<td>0</td>
<td>0.01</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>568</td>
<td>II+</td>
<td>0.01</td>
<td>0</td>
<td>0.01</td>
<td>0.32</td>
<td>0.29</td>
<td>0.34</td>
<td>0.02</td>
<td>0.01</td>
<td>0.01</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1136</td>
<td>II++</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.22</td>
<td>0.15</td>
<td>0.52</td>
<td>0.02</td>
<td>0.01</td>
<td>0.06</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>291</td>
<td>III-</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.05</td>
<td>0.02</td>
<td>0.03</td>
<td>0.29</td>
<td>0.23</td>
<td>0.36</td>
<td>0.01</td>
<td>0</td>
<td>0.02</td>
</tr>
<tr>
<td>206</td>
<td>III+</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.01</td>
<td>0.01</td>
<td>0</td>
<td>0.29</td>
<td>0.16</td>
<td>0.45</td>
<td>0.01</td>
<td>0.02</td>
<td>0.06</td>
</tr>
<tr>
<td>569</td>
<td>III++</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.19</td>
<td>0.15</td>
<td>0.53</td>
<td>0.02</td>
<td>0.02</td>
<td>0.08</td>
</tr>
<tr>
<td>216</td>
<td>IV-</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.07</td>
<td>0.01</td>
<td>0.01</td>
<td>0.32</td>
<td>0.26</td>
<td>0.34</td>
</tr>
<tr>
<td>190</td>
<td>IV+</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.01</td>
<td>0.26</td>
<td>0.33</td>
<td>0.41</td>
</tr>
<tr>
<td>328</td>
<td>IV++</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.27</td>
<td>0.26</td>
<td>0.46</td>
</tr>
</tbody>
</table>

Table 2: transition matrix

In the transition matrix the first column indicates the number of observations for every state transition.

---

\(^{13}\)In this matrix and in those in Appendix C rows may not sum to one due to rounding.
4.2 Results

The number of observations is not equally distributed among states, but every state appears to have a sufficient number of observations. The ergodic distribution is reported in Table 3.

<table>
<thead>
<tr>
<th></th>
<th>I-</th>
<th>I+</th>
<th>I++</th>
<th>II-</th>
<th>II+</th>
<th>III-</th>
<th>III+</th>
<th>IV-</th>
<th>IV+</th>
<th>IV++</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.04</td>
<td>0.03</td>
<td>0.05</td>
<td>0.06</td>
<td>0.04</td>
<td>0.09</td>
<td>0.19</td>
<td>0.19</td>
</tr>
</tbody>
</table>

Table 3: ergodic distribution

Table 4 reports the ergodic distribution with respect to a normalization of the distribution’s mass in every GDP class. This representation of the ergodic distribution highlights the growth rate dynamics within each GDP class in steady state.

<table>
<thead>
<tr>
<th></th>
<th>-</th>
<th>+</th>
<th>++</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>0.39</td>
<td>0.27</td>
<td>0.34</td>
</tr>
<tr>
<td>II</td>
<td>0.35</td>
<td>0.22</td>
<td>0.42</td>
</tr>
<tr>
<td>III</td>
<td>0.32</td>
<td>0.19</td>
<td>0.49</td>
</tr>
<tr>
<td>IV</td>
<td>0.29</td>
<td>0.28</td>
<td>0.43</td>
</tr>
</tbody>
</table>

Table 4: ergodic distribution normalized for each GDP class

Finally, in Table 5 we report the cross-country income distribution of the first and last year, along with the ergodic distribution in terms of GDP classes only.

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>1950</td>
<td>0.31</td>
<td>0.53</td>
<td>0.13</td>
<td>0.02</td>
</tr>
<tr>
<td>1998</td>
<td>0.16</td>
<td>0.38</td>
<td>0.22</td>
<td>0.25</td>
</tr>
<tr>
<td>Ergodic</td>
<td>0.03</td>
<td>0.13</td>
<td>0.19</td>
<td>0.65</td>
</tr>
</tbody>
</table>

Table 5: first and last year distribution vs ergodic distribution

The ergodic distribution represents the long-run, or invariant, distribution. Its existence is generally guaranteed if the process is irreducible, aperiodic and positive persistent. In our case these properties are satisfied. The ergodic distribution can also include some states with zero mass when there exists only one irreducible closed set of positive persistent aperiodic states, and the remaining states are transient (see Isacson and Madsen (1976), p. 74).
4.2 Results

In Figure 3 we report the contour plots of kernel density estimations for 3-year transitions of the growth rate within every GDP class (see Durlauf and Quah (1999) for details on stochastic kernels).

The vertical and horizontal axis respectively refer to year $t$ and year $t + 3$. We superimpose a grid representing our growth rate classes and a 45° line, which helps to identify the probabilities of acceleration or deceleration of growth rates. This technique complements the estimate of the transition matrix avoiding the problem of the discretization of the growth rates’ space, which may introduce spurious dynamic ef-
4.2 Results

Nonlinearities in the growth process

For each GDP class we assess whether the predictions from the theoretical model find support.

In GDP class I we expect a deceleration and/or a stagnation of growth. With respect to the other GDP classes we observe that, given a high growth rate, the probability of a low growth rate is the highest ($\hat{p}_{I++,} = 0.28$ vs $\hat{p}_{II++,} = 0.24$, $\hat{p}_{III++,} = 0.21$ and $\hat{p}_{IV++,} = 0.27$, where $\hat{p}_{r,w} = \sum_{q \in \{I,...,IV\}} \hat{p}_{ir,qu}^q$ for $i \in \{I,...,IV\}$ and for $r,w \in \{-,+,++,\}$)\(^{[16]}\) and the probability of another high growth rate is the lowest after that of GDP class IV ($\hat{p}_{I++,} = 0.51$ vs $\hat{p}_{II++,} = 0.58$, $\hat{p}_{III++,} = 0.61$ and $\hat{p}_{IV++,} = 0.46$)\(^{[17]}\). In addition, given a low growth rate, the probability of another low growth rate is greatest in GDP class I, in particular with respect to GDP classes III and IV ($\hat{p}_{1,-} = 0.49$ vs $\hat{p}_{II,-} = 0.47$, $\hat{p}_{III,-} = 0.35$ and $\hat{p}_{IV,-} = 0.39$)\(^{[18]}\). Finally, given a medium growth rate, the probability of a medium growth rate is the highest ($\hat{p}_{I+,-} = 0.48$ vs $\hat{p}_{II+,-} = 0.30$, $\hat{p}_{III+,-} = 0.19$ and $\hat{p}_{IV+,-} = 0.33$)\(^{[19]}\). The observed persistence at medium and low growth rates should imply that, in the long run, countries in GDP class I should spend a relevant amount of time in this growth rate classes. This insight finds a confirmation in Table\(^{[4]}\) first row, where 66% of the mass is in the first two growth rate classes. In Appendix\(^{[C]}\) we report the estimates of transition matrices with 1-year lags and with 3-year average growth rates, showing similar results. Therefore, our findings do not seem to depend on the possible autocorrelation of shocks.

Furthermore, Figure\(^{[3]}\) confirms that in GDP class I having a transition to a medium growth rate is the most likely event for almost any initial growth rate, since the peaks

---

\(^{[15]}\) See Durlauf et al. (2004), pp. 57 - 58, for a discussion of this problem. Below we consider the case of a continuous GDP space.

\(^{[16]}\) These “modified” transition probabilities simply refer to each (stochastic) submatrix corresponding to an income level. E.g. the value $\hat{p}_{1++,} = 0.28$ is obtained by summing $\hat{p}_{I++,} = 0.22$ and $\hat{p}_{II++,} = 0.06$.

\(^{[17]}\) Tests of equality between $\hat{p}_{I++,}$ and, respectively, $\hat{p}_{II++,}$, $\hat{p}_{III++,}$ and $\hat{p}_{IV++,}$ return the following p-values: 0.06, 0.01, and 0.38. Tests of equality between $\hat{p}_{I++}$ and, respectively, $\hat{p}_{II++}$, $\hat{p}_{III++}$ and $\hat{p}_{IV++}$ return the following p-values: 0.01, 0, and 0.09 (see Appendix\(^{[E]}\) for more details on these tests).

\(^{[18]}\) Tests of equality between $\hat{p}_{1-}$ and, respectively, $\hat{p}_{II-}$, $\hat{p}_{III-}$ and $\hat{p}_{IV-}$ return the following p-values: 0.37, 0, and 0.01. Clearly, we cannot reject the hypothesis that $\hat{p}_{1-}$ is equal to $\hat{p}_{II-}$.

\(^{[19]}\) In this case the hypothesis of equality between $\hat{p}_{1+}$ and the other probabilities can be rejected at 0 level of significance.
of the stochastic kernel are located in correspondence of medium growth rates. This finding provides some empirical support to trajectory A in Figure 2. The comparison with stochastic kernels of the other GDP classes highlights the relatively low persistence at high growth rates (in particular with respect to GDP classes II and III), and the relatively high persistence at low growth rates, in particular with respect to GDP classes III and IV.

In GDP class II we should observe the beginning of the acceleration phase. In accordance with Figure 2 persistence at low growth rates is similar to GDP class I (\(\hat{p}_{II-,\cdot} = 0.48\) vs \(\hat{p}_{I-,\cdot} = 0.49\)). Persistence at medium growth rates is similar to GDP class IV (\(\hat{p}_{II+,\cdot} = 0.30\) vs \(\hat{p}_{IV+,\cdot} = 0.33\)), which is in accordance with Figure 2 with respect to the existence of Trajectory A. However, with respect to GDP class I, the probability of transition from medium to high growth rate is much higher (\(\hat{p}_{II+,\cdot} = 0.35\) vs \(\hat{p}_{I+,\cdot} = 0.24\)), as well as the probability of persisting at high growth rate (\(\hat{p}_{II+,\cdot} = 0.58\) vs \(\hat{p}_{I+,\cdot} = 0.51\)). Note also that the mass of probability of high growth rates in the ergodic distribution reported in Table 4 increases from 0.34 to 0.42. Again, transition matrices with 1-year lags and with 3-year average growth rates show similar results. Figure 3 corroborates these findings. In particular, with respect to GDP class I the “ridge” of the stochastic kernel in GDP class II rotates clockwise and therefore the peaks appears to be placed further on the right, especially for the medium/high growth rate.

GDP class III should be characterized by both (i) acceleration of growth and (ii) persistence at high growth rates. In the transition matrix the set of relevant probabilities for point (i) is given by: \(\hat{p}_{I+,\cdot} \hat{p}_{II+,\cdot} \hat{p}_{III+,\cdot} \hat{p}_{IV+,\cdot}\) and \(\hat{p}_{IV+,\cdot}\). The estimated values are, respectively: \(\hat{p}_{I+,\cdot} = 0.24\), \(\hat{p}_{II+,\cdot} = 0.35\), \(\hat{p}_{III+,\cdot} = 0.51\) and \(\hat{p}_{IV+,\cdot} = 0.42\). Thus, it appears that a country in GDP class III is relatively more likely to show accelerating growth, in accordance with the prediction of the model in Figure 2. More precisely, we find that the probability to increase an already sustained growth rate rises with income for the first three GDP classes, and then decreases in the fourth. As regards point (ii) note that, for a country with a high growth rate, the probability of maintaining such rate is highest in income class III. The relevant value is \(\hat{p}_{III+,\cdot} = 0.61\) against \(\hat{p}_{I+,\cdot} = 0.51\), \(\hat{p}_{II+,\cdot} = 0.58\) and \(\hat{p}_{IV+,\cdot} = 0.46\). Note also that the probability of persistence of low growth rates is the smallest (\(\hat{p}_{III-,\cdot} = 0.35\) vs \(\hat{p}_{I-,\cdot} = 0.49\),

\[\text{Tests of equality of \(p_{III+,\cdot}\) and \(p_{I+,\cdot}\) give the following p-values: 0 and 0.01.}\]

\[\text{Tests of equality between \(p_{III+,\cdot}\) and \(p_{II+,\cdot}\) and \(p_{IV+,\cdot}\) return the following p-values: 0, 0 and 0.04.}\]

\[\text{The hypothesis of equality between \(p_{III+,\cdot}\) and \(p_{I+,\cdot}\) and \(p_{IV+,\cdot}\) is strongly rejected. Instead the hypothesis of equality between \(p_{III+,\cdot}\) and \(p_{I+,\cdot}\) gives a p-value of 0.12.}\]
4.2 Results 4 EMPIRICAL ANALYSIS

\( \hat{p}_{II--} = 0.48 \) and \( \hat{p}_{IV--} = 0.39 \)\(^{23}\) as well as the probability of persistence at medium growth rates (\( \hat{p}_{III+,+} = 0.19 \) vs \( \hat{p}_{I+,+} = 0.48 \), \( \hat{p}_{II+,+} = 0.30 \) and \( \hat{p}_{IV+,+} = 0.33 \))\(^{24}\) There is a further increase in the probability of high growth rates in the ergodic distribution reported in Table \(4\) from 0.42 to 0.49\(^{25}\). Again, estimates of transition matrices with 1-year lag and with 3-year average growth rates show similar results (see Appendix \(C\)). In Figure \(3\) we observe a rightward shift of the ridge of the kernel. Now the peak of stochastic kernel for medium/high growth rates is more clearly in the high growth rate class.

In GDP class \(IV\) deceleration from high growth to medium growth becomes a more likely event (\( \hat{p}_{IV++,+} = 0.26 \), \( \hat{p}_{III++,+} = 0.17 \), \( \hat{p}_{II++,+} = 0.16 \) and \( \hat{p}_{I++,+} = 0.21 \))\(^{26}\). In this GDP class there is a relatively high persistence at medium growth rates, in particular with respect to GDP class \(III\) (\( \hat{p}_{IV++,+} = 0.33 \), \( \hat{p}_{III++,+} = 0.19 \))\(^{27}\). The ergodic distribution for this GDP class in Table \(4\) shows the highest value of the probability mass for medium growth rate is in GDP class \(IV\) (0.28 vs 0.27, 0.22, 0.19). Estimates with 3-year average growth rates reported in Appendix \(C\) which should reduce at the minimum the possible presence of autocorrelation, show that in GDP class \(IV\) the probability of medium growth rate is the highest (0.42 vs 0.23 and 0.36)\(^{28}\) with respect to the other growth rate classes (it is also the highest with respect to the other GDP classes (0.42 vs 0.36, 0.30 and 0.21)\(^{29}\). The shape of the stochastic kernel in GDP class \(IV\) reveals that, indeed, in this class having a medium growth rate is the most likely event starting from any level of the growth rate.

**Existence of poverty traps** A key question for growth empirics is the existence of poverty traps. The ergodic distribution in terms of only GDP classes in Table \(5\) shows that the proportion of countries in GDP classes \(I\) and \(II\) strongly tends to decrease in favour of GDP class \(IV\), as is the case in absence of a poverty trap in GDP class \(I\). However, the comparison between the initial, final and ergodic distributions reveals that convergence is very slow.

Table \(5\) shows that full convergence was far from being achieved in 1998. Follow-

\(^{23}\) All hypotheses of equality are strongly rejected with the exception of \( \hat{p}_{IV--}, \) which gives a p-value of 0.12.

\(^{24}\) All tests of equality return a p-value of 0.

\(^{25}\) A test of equality gives a p-value of 0.04.

\(^{26}\) Also in this case the hypothesis of equality between \( \hat{p}_{IV++,+} \) and the other probabilities can be rejected. Respectively, the tests give the following p-values: 0, 0 and 0.02.

\(^{27}\) As noted, this difference is statistically significant.

\(^{28}\) Tests of equality between the first and the other values return the following p-values: 0 and 0.16.

\(^{29}\) Tests of equality between the first and the other values return the following p-values: 0.01, 0 and 0.
4.2 Results

ing Shorrocks (1978), we compute the asymptotic half life of this process. It equals about 18.41 periods (i.e. 55 years), so that at least 110 years from the end of our observation’s period should be necessary to have full convergence.\(^\text{20}\) A first remark regards the stability of the process in the long run. We estimate a transition matrix from observations spanning 48 years, while convergence would occur in more than twice those years; therefore the significance of the ergodic distribution can be questionable.

As noted in the introduction, if all countries follow a common trend, we cannot be sure that the behaviour we identify, e.g. for GDP class III \([8.29, 9.21]\) (which corresponds to \([4000, 10000]\)) in the period 1950-1998, may be observed in the same GDP class say in 2050-2098. However, the ergodic distribution can nonetheless provide some insights on the long-run tendency of the cross-country distribution. A slow adjustment process may mean for some countries a long period of very low growth, as it may be the case of African countries. Moreover, even if the distribution dynamics shows a reduction in the weight of GDP classes \(I-III\) in favour of \(IV\), in the ergodic distribution a relevant part of mass \((0.35)\) is contained in the first three GDP classes. This implies that there may always exist a set of countries remaining poor in absolute terms, even if we do not find unambiguous support of a poverty trap in GDP class \(I\).\(^\text{31}\)

The fact that in the ergodic distribution there remains a positive fraction of countries in the first three GDP classes depends on the significant transition probabilities of moving from a GDP class to a class with a lower GDP. This is clearly observable from the transition matrix for GDP classes in Table 6 and Table 7, which refer respectively to 3-year and 15-year transitions.\(^\text{32}\)

\(^{20}\)The asymptotic half life is defined as \(h = -\log 2 / \log |\lambda_2|\), where \(\lambda_2\) is the second largest eigenvalue of the transition matrix. In our case \(\lambda_2 \approx 0.9631\). This measure of the speed of convergence is based on the time the process takes from period \(t\) to reach half of the distance from its equilibrium level (the ergodic distribution).

\(^{31}\)Obviously, the members of this set can change over time, since the ergodicity of process means that every country has a positive probability to visit each state.

\(^{32}\)Tests on the lower off-diagonal elements show that they are statistically different from zero. Similar values are found in the transition matrices with 1-year lags and 3-year average growth rates.
4.2 Results

Table 6: transition matrix for GDP classes: 3-year lag

<table>
<thead>
<tr>
<th>Obs.</th>
<th>States</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>1116</td>
<td>I</td>
<td>0.89</td>
<td>0.11</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2574</td>
<td>II</td>
<td>0.03</td>
<td>0.92</td>
<td>0.05</td>
<td>0</td>
</tr>
<tr>
<td>1066</td>
<td>III</td>
<td>0</td>
<td>0.03</td>
<td>0.88</td>
<td>0.09</td>
</tr>
<tr>
<td>734</td>
<td>IV</td>
<td>0</td>
<td>0</td>
<td>0.02</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>Ergodic</td>
<td>0.03</td>
<td>0.13</td>
<td>0.19</td>
<td>0.65</td>
</tr>
</tbody>
</table>

Table 7: transition matrix for GDP classes: 15-year lag

<table>
<thead>
<tr>
<th>Obs.</th>
<th>States</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>866</td>
<td>I</td>
<td>0.61</td>
<td>0.39</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1192</td>
<td>II</td>
<td>0.07</td>
<td>0.70</td>
<td>0.22</td>
<td>0.01</td>
</tr>
<tr>
<td>736</td>
<td>III</td>
<td>0</td>
<td>0.05</td>
<td>0.46</td>
<td>0.48</td>
</tr>
<tr>
<td>432</td>
<td>IV</td>
<td>0</td>
<td>0</td>
<td>0.11</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td>Ergodic</td>
<td>0.01</td>
<td>0.04</td>
<td>0.18</td>
<td>0.77</td>
</tr>
</tbody>
</table>

We also consider long transitions because, as remarked by Kremer et al. (2001), p. 284, this should strongly reduce the effect of short-run fluctuations, and therefore may be more appropriate to study poverty traps, a typical long-run phenomenon. With 15-year transitions, as predictable, there is much less persistence in every state as reflected by the decrease of the elements on the principal diagonal. In both cases transitions from a higher to a lower GDP class do not appear as a negligible aspect, as also reflected in the ergodic distribution where the proportion of countries in GDP class IV increases but is still relatively lower than 1. This is particularly remarkable because we are considering absolute levels of GDP.

As the definition of a discrete state space may affect the probabilistic nature of data, in particular for what concerns the estimation of long-run tendencies, we evaluate the distribution dynamics with a continuous state space, in particular by comparing the initial, final and ergodic density of the countries.

Transitions from GDP class IV to GDP class III regards the following countries: Gabon, Venezuela, Trinidad & Tobago, Kuwait, Qatar and Saudi Arabia. These countries are excluded from the analysis by Kremer et al. (2001), as their GDP depends for more than 15% on oil or other non-renewable resources. As a consequence, they find that the highest GDP class (relative in their case) is a quasi-absorbing state. We question this selection criterium as it eliminates from the analysis a fundamental aspect such as the structure of the economy. We conjecture that the mentioned countries make such transitions because their aggregate GDP is more unstable, being heavily based on the output of highly volatile sectors.

See, e.g. Durlauf et al. (2004), pp. 57 - 58.

To estimate the ergodic distribution we follow Johnson (2005). The author kindly provided us with the procedure, now available at http://irving.vassar.edu/faculty/pj/pj.htm.
Figures 4 and 5 confirm that the probability mass in the long run is not completely concentrated at the highest GDP levels, especially with 3-year transitions. This result is in contrast with the dynamics predicted by technological diffusion models. For instance in Lucas (2000), once a country leaves stagnation, it can just proceed towards higher income levels.

The overall results support the dynamics depicted in Section 3, i.e. the dynamics and the distribution of the probability masses looks coherent with Figure 2.

5 Conclusions

The main result of the paper is the detection of nonlinearities in the growth process. In particular, we find support to the picture of a range of decreasing or persistently low growth rates, followed by a phase of accelerating growth (take-off), which eventually decelerates once a country reaches a certain level of per capita GDP. However, this process appears rather slow, which may mean a long period of low or zero growth for a number of low-income countries. This is coupled to the lack of a tendency for all countries to reach, and to remain within, the highest GDP class due to a non-negligible probability of “reversal of fortune”.

In general, contributions analysing cross-country growth dynamics should take into account the nonlinear pattern of the growth rate. In this respect, this paper provides some “stylized facts” which a model aiming to reproduce the development path of a country should match. In particular, if technological diffusion is at the heart of
economic growth across countries, then its features should differ from those presented in Lucas (2000), for instance by allowing for a variable adoption speed of new technologies and reversibility of adoption\footnote{36} as well as for the consideration of the structure of the economy.

Finally, we believe that traditional development theories still provide interesting insights on the emergence of nonlinearities, as the focus on structural change, that modern growth literature only recently started to explore intensively (see Galor and Weil (2000)). This is the direction of our current research.

\footnote{36}By reversibility of adoption we mean the possibility that a country loses the capability of exploiting technologies previously used, due for instance to the depletion of the human capital stock (e.g. for starvation, epidemics, wars, etc.). This would follow the insights of Basu and Weil (1998) on appropriate technologies.
References


REFERENCES


## A Country List

<table>
<thead>
<tr>
<th>Country</th>
<th>Rank</th>
<th>Country</th>
<th>Rank</th>
<th>Country</th>
<th>Rank</th>
<th>Country</th>
<th>Rank</th>
<th>Country</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFRICA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Algeria</td>
<td>1</td>
<td>Angola</td>
<td>2</td>
<td>Benin</td>
<td>3</td>
<td>Botswana</td>
<td>4</td>
<td>Cameroon</td>
<td>5</td>
</tr>
<tr>
<td>Congo</td>
<td>10</td>
<td>Côte d’Ivoire</td>
<td>11</td>
<td>Djibouti</td>
<td>12</td>
<td>Egypt</td>
<td>13</td>
<td>Gabon</td>
<td>14</td>
</tr>
<tr>
<td>Gambia</td>
<td>15</td>
<td>Ghana</td>
<td>16</td>
<td>Kenya</td>
<td>17</td>
<td>Liberia</td>
<td>18</td>
<td>Madagascar</td>
<td>19</td>
</tr>
<tr>
<td>Mali</td>
<td>20</td>
<td>Mauritania</td>
<td>21</td>
<td>Mauritius</td>
<td>22</td>
<td>Morocco</td>
<td>23</td>
<td>Mozambique</td>
<td>24</td>
</tr>
<tr>
<td>Namibia</td>
<td>25</td>
<td>Niger</td>
<td>26</td>
<td>Reunion</td>
<td>28</td>
<td>Rwanda</td>
<td>29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Senegal</td>
<td>30</td>
<td>Seychelles</td>
<td>31</td>
<td>Sierra Leone</td>
<td>32</td>
<td>Somalia</td>
<td>33</td>
<td>South Africa</td>
<td>34</td>
</tr>
<tr>
<td>Sudan</td>
<td>35</td>
<td>Swaziland</td>
<td>36</td>
<td>Tanzania</td>
<td>37</td>
<td>Togo</td>
<td>38</td>
<td>Tunisia</td>
<td>39</td>
</tr>
<tr>
<td>Uganda</td>
<td>40</td>
<td>Zambia</td>
<td>41</td>
<td>Zimbabwe</td>
<td>42</td>
<td>LATIN AMERICA</td>
<td>43</td>
<td>Argentina</td>
<td></td>
</tr>
<tr>
<td>Brazil</td>
<td>44</td>
<td>Chile</td>
<td>45</td>
<td>Colombia</td>
<td>46</td>
<td>Mexico</td>
<td>47</td>
<td>Peru</td>
<td>48</td>
</tr>
<tr>
<td>Uruguay</td>
<td>49</td>
<td>Venezuela</td>
<td>50</td>
<td>Bolivia</td>
<td>51</td>
<td>Costa Rica</td>
<td>52</td>
<td>Cuba</td>
<td>53</td>
</tr>
<tr>
<td>Dominican Rep.</td>
<td>54</td>
<td>Ecuador</td>
<td>55</td>
<td>El Salvador</td>
<td>56</td>
<td>Guatemala</td>
<td>57</td>
<td>Haiti</td>
<td>58</td>
</tr>
<tr>
<td>Honduras</td>
<td>59</td>
<td>Jamaica</td>
<td>60</td>
<td>Nicaragua</td>
<td>61</td>
<td>Panama</td>
<td>62</td>
<td>Paraguay</td>
<td>63</td>
</tr>
<tr>
<td>Puerto Rico</td>
<td>64</td>
<td>Trin. Tobago</td>
<td>65</td>
<td>OFF WESTERN</td>
<td>66</td>
<td>Australia</td>
<td>67</td>
<td>New Zealand</td>
<td></td>
</tr>
<tr>
<td>Canada</td>
<td>68</td>
<td>United States</td>
<td>69</td>
<td>WEST ASIA</td>
<td>70</td>
<td>Bahrain</td>
<td>71</td>
<td>Iran</td>
<td></td>
</tr>
<tr>
<td>Iraq</td>
<td>72</td>
<td>Israel</td>
<td>73</td>
<td>Jordan</td>
<td>74</td>
<td>Kuwait</td>
<td>75</td>
<td>Lebanon</td>
<td>76</td>
</tr>
<tr>
<td>Oman</td>
<td>77</td>
<td>Qatar</td>
<td>78</td>
<td>Saudi Arabia</td>
<td>79</td>
<td>Syria</td>
<td>80</td>
<td>Turkey</td>
<td>81</td>
</tr>
<tr>
<td>UAE</td>
<td>82</td>
<td>Yemen</td>
<td>83</td>
<td>W.Bank Gaza</td>
<td>84</td>
<td>EAST ASIA</td>
<td>85</td>
<td>China</td>
<td></td>
</tr>
<tr>
<td>India</td>
<td>86</td>
<td>Indonesia</td>
<td>87</td>
<td>Japan</td>
<td>88</td>
<td>Philippines</td>
<td>89</td>
<td>South Korea</td>
<td>90</td>
</tr>
<tr>
<td>Thailand</td>
<td>91</td>
<td>Taiwan</td>
<td>92</td>
<td>Bangladesh</td>
<td>93</td>
<td>Burma</td>
<td>94</td>
<td>Hong Kong</td>
<td>95</td>
</tr>
<tr>
<td>Malaysia</td>
<td>96</td>
<td>Nepal</td>
<td>97</td>
<td>Pakistan</td>
<td>98</td>
<td>Singapore</td>
<td>99</td>
<td>Sri Lanka</td>
<td>100</td>
</tr>
<tr>
<td>Afghanistan</td>
<td>101</td>
<td>Cambodia</td>
<td>102</td>
<td>Laos</td>
<td>103</td>
<td>Mongolia</td>
<td>104</td>
<td>North Korea</td>
<td>105</td>
</tr>
<tr>
<td>Vietnam</td>
<td>106</td>
<td>EUROPE</td>
<td>107</td>
<td>Austria</td>
<td>108</td>
<td>Belgium</td>
<td>109</td>
<td>Denmark</td>
<td></td>
</tr>
<tr>
<td>Finland</td>
<td>110</td>
<td>France</td>
<td>111</td>
<td>Germany</td>
<td>112</td>
<td>Italy</td>
<td>113</td>
<td>Netherlands</td>
<td>114</td>
</tr>
<tr>
<td>Norway</td>
<td>115</td>
<td>Sweden</td>
<td>116</td>
<td>Switzerland</td>
<td>117</td>
<td>UK</td>
<td>118</td>
<td>Ireland</td>
<td>119</td>
</tr>
<tr>
<td>Greece</td>
<td>120</td>
<td>Portugal</td>
<td>121</td>
<td>Spain</td>
<td>122</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 8: country list

## B Comparing methods to identify growth regimes

In this appendix we compare the method proposed in this paper to identify growth regimes, based on a nonparametric estimation of the relation between the growth rate and the level of GDP, and the methods to identify thresholds values in explanatory...
variables proposed by Breiman et al. (1984) and Hansen (2000). The former has been applied to the study of economic growth by Durlauf and Johnson (1995), and consists in the sequential application of a splitting rule to the data according to their value, until the “best” partition (a tree) is found. The latter is a procedure to identify possible statistically significant thresholds in linear regressions by a likelihood ratio statistics.

With respect to the procedure followed here, we are faced with a trade-off. If growth regimes are characterized by strong nonlinearities in the relation between growth rates and GDP levels, then a nonparametric estimate is preferable because it is truly local, and therefore more capable to keep track of all possible nonlinearities. Moreover, it identifies a continuous path, by an appropriate smoothing. The drawback is that the exact levels of the thresholds we use to build the GDP classes cannot be identified in a statistically rigorous way. For instance, when looking for a threshold separating the first growth regime from the second, we choose the GDP level where the estimated path reaches a local minimum, as suggested by our theoretical framework. However, we do not have a confidence interval around this level. In addition, the localization of this point depends on the fact that, as noted, the nonparametric estimation necessarily returns a continuous path while the change in regime could be actually marked by a discontinuity.

On the other hand, the methods of Breiman et al. (1984) and Hansen (2000) identify the thresholds more rigorously as they find GDP thresholds, respectively, with some consistency properties or a confidence band. These GDP thresholds refer to GDP classes in which the relation between the growth rate and GDP is different, and may correspond to discontinuities in the growth path. The drawback is that both procedures, while searching for thresholds, assume that the different relations are essentially linear.

For instance, the classification method of Breiman et al. (1984) assumes that the different groups of observations on GDP, are simply different in terms of the average growth rate. Hence, this method does not take into account that growth regimes, at least as they are defined by the theoretical model we wish to test, can themselves be characterized by nonlinearities.

Therefore, there do not appear specific reasons for unambiguously prefer one method to the other on theoretical grounds. From an empirical point of view we remark that, for the type of data we are examining, a nonparametric estimate has an advantage in that it does not suffer from the relatively high dispersion of data and the likely presence
of heteroskedasticity.

### B.1 Regression tree

When we apply the Breiman et al. (1984) algorithm to our data, we find that a tree with three terminal nodes represents the best partition as it minimizes the cross-validated error, basically corresponding to a sum of squared residuals. Data on GDP would be partitioned in three classes: $(-\infty, 7.34)$, $[7.34, 9.98)$, and $[9.98, \infty)$. However, if we follow the 1-SE rule we would choose the case in which all observations belong to one group (a tree with no branches). The only information on the relation between GDP and the growth rate in the three classes is that the average growth rates associated to the observations on GDP they contain are respectively: 0.013, 0.023 and 0.010. We conjecture that the high dispersion in the data and the presence of only one explanatory variable are responsible for the difficulty in identifying a partition. At any rate, note that on average the relation between growth rates and GDP follows our predicted path, being first increasing and then decreasing.

In Figures 6 and 7, following Durlauf and Johnson (1995), we run a simple linear regression between growth rate and GDP in, respectively, the three GDP classes identified by the regression tree analysis and the three GDP classes suggested by the non-parametric regression. We define our regimes on the basis of the minimum reached by the estimate in Figure 1 at 6.84 and the maximum at 8.85. We consider three regimes instead of four to focus on the changes in the slope of the growth rate-GDP relation. This makes the comparison with the other methods simpler.

---

38We used the package `rpart`, based on Breiman et al. (1984). See Therneau and Atkinson (2005).

39The 1-SE rule is a method to choose among the error-minimizing trees, where (cross-validated) error minimization is done with respect to different levels of the cost-complexity parameter. The latter is a parameter which penalizes the complexity of the data partition, that is the number of terminal nodes of a regression tree. The 1-SE rule suggests to choose the smallest tree which has a cross-validated error within one standard error from the minimum estimated error. See Breiman et al. (1984), p. 237, for more details.

40In order to reduce the dispersion in the data, we considered 3-year and 5-year average growth rates, without obtaining substantial differences. Smoothing the data in this case has the disadvantage of reducing the amount of curvature present in the data, and therefore makes what we consider the determinant of growth regimes less capable of guiding the data partition.

41For clarity, the observations are omitted.
We note that in Figure 6 the range of GDP in which the growth rate is increasing in income is not identified. Moreover, only the negative coefficient of the relation in the first GDP class is statistically significant. Finally, the growth path appears to be strongly discontinuous. Differently, the path in Figure 7 identifies an intermediate range of GDP where the growth rate is increasing. All slope coefficients are statistically significant and the growth path appears to be essentially continuous (note the confidence bands). Overall, the results in Figure 7 are consistent with the nonparametric estimate and with the results on the distribution dynamics.

B.2 Threshold regression model

Hansen (2000) proposes the following procedure. In the first step we perform a LM (Lagrange multiplier) test on the full sample, where the null hypothesis is the nonexistence of threshold effects. If this null is rejected, we can conclude that there exist different regimes, and a LR (likelihood ratio) test indicates the confidence range of GDP including the possible threshold. The next step consists in testing whether other growth regimes are present within the two regimes previously identified. If these tests are not passed, we apply again a LR test to identify other thresholds. The procedure stops when the hypothesis of no threshold effects is not rejected in all regimes. Table 9 reports the results of this recursive procedure for our sample.

42 All calculations are made by a GAUSS code available on Hansen’s website.
In the full sample we reject the null hypothesis of no threshold effects (p-value is equal to 0). This is indicated by “NO” in the first row. The LR test at 5% significance level indicates a possible threshold in the range [8.53, 8.84]. Following Hansen (2000) we take 8.60, the minimum reached by the LR statistics, as a threshold and run a LM test in each of the two resulting growth regimes (this is the second step). In the regime with GDP > 8.60 we found no other regimes (the p-value is equal to 0.44), while in the regime with GDP ≤ 8.60 we found a threshold effect (p-value is equal to 0). The LR test at 5% significance level indicates a possible threshold in the range [6.09, 7.38]. We take 7.34, the minimum reached by the LR statistics, and run a LM test in each of the two resulting growth regimes (third step). Both LM tests do not reject the null (the p-values are equal to 0.29 and 0.87) and the procedure stops.

Therefore, we find three growth regimes, and two thresholds in the ranges [6.09, 7.38] and [8.53, 8.84]. We observe that these two ranges are compatible with the results of the nonparametric regression, where the threshold values we selected are 6.84 and 8.85. Note that the first range is particularly large, an aspect which has a counterpart in the relatively high confidence band at low GDP levels in the nonparametric estimate.

Finally, we remark that Hansen (2000) does not provide any further indication on how to choose an exact value of the threshold within the estimated ranges.

We conclude that, for this type of data, the nonparametric regression is an appropriate tool to identify growth regimes.

### C Other estimates

In this appendix we report alternative estimates. We do not provide any discussion of the results, except that they support our previous findings.
C.1 Estimates with 1-year lags

In this section we present the transition matrix for 1-year transitions, i.e. from \((y_t, g_t)\) to \((y_{t+1}, g_{t+1})\), along with the tables for the ergodic distribution and the distribution dynamics.

<table>
<thead>
<tr>
<th>N. Obs</th>
<th>States</th>
<th>I-</th>
<th>I+</th>
<th>I++</th>
<th>II-</th>
<th>II+</th>
<th>II++</th>
<th>III-</th>
<th>III+</th>
<th>III++</th>
<th>IV-</th>
<th>IV+</th>
<th>IV++</th>
</tr>
</thead>
<tbody>
<tr>
<td>406</td>
<td>I-</td>
<td>0.52</td>
<td>0.16</td>
<td>0.32</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>347</td>
<td>I+</td>
<td>0.22</td>
<td>0.56</td>
<td>0.21</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>401</td>
<td>I++</td>
<td>0.23</td>
<td>0.20</td>
<td>0.45</td>
<td>0.04</td>
<td>0.02</td>
<td>0.07</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>896</td>
<td>II-</td>
<td>0.03</td>
<td>0</td>
<td>0.01</td>
<td>0.49</td>
<td>0.18</td>
<td>0.28</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>603</td>
<td>II+</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.28</td>
<td>0.41</td>
<td>0.31</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1171</td>
<td>II++</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.22</td>
<td>0.15</td>
<td>0.59</td>
<td>0.01</td>
<td>0.01</td>
<td>0.03</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>304</td>
<td>III-</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.03</td>
<td>0.01</td>
<td>0.01</td>
<td>0.47</td>
<td>0.17</td>
<td>0.30</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>215</td>
<td>III+</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.30</td>
<td>0.26</td>
<td>0.42</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>597</td>
<td>III++</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.16</td>
<td>0.16</td>
<td>0.63</td>
<td>0.01</td>
<td>0.01</td>
<td>0.04</td>
</tr>
<tr>
<td>220</td>
<td>IV-</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.02</td>
<td>0.01</td>
<td>0.01</td>
<td>0.48</td>
<td>0.24</td>
<td>0.23</td>
</tr>
<tr>
<td>215</td>
<td>IV+</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.27</td>
<td>0.33</td>
<td>0.40</td>
<td></td>
</tr>
<tr>
<td>359</td>
<td>IV++</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.16</td>
<td>0.26</td>
<td>0.58</td>
<td></td>
</tr>
</tbody>
</table>

Table 10: transition matrix

<table>
<thead>
<tr>
<th></th>
<th>I-</th>
<th>I+</th>
<th>I++</th>
<th>II-</th>
<th>II+</th>
<th>II++</th>
<th>III-</th>
<th>III+</th>
<th>III++</th>
<th>IV-</th>
<th>IV+</th>
<th>IV++</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>0.02</td>
<td>0.02</td>
<td>0.06</td>
<td>0.04</td>
<td>0.08</td>
<td>0.06</td>
<td>0.04</td>
<td>0.11</td>
<td>0.15</td>
<td>0.15</td>
<td>0.24</td>
<td></td>
</tr>
</tbody>
</table>

Table 11: ergodic distribution

<table>
<thead>
<tr>
<th>States</th>
<th>-</th>
<th>+</th>
<th>++</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>0.37</td>
<td>0.29</td>
<td>0.34</td>
</tr>
<tr>
<td>II</td>
<td>0.34</td>
<td>0.23</td>
<td>0.43</td>
</tr>
<tr>
<td>III</td>
<td>0.30</td>
<td>0.20</td>
<td>0.51</td>
</tr>
<tr>
<td>IV</td>
<td>0.28</td>
<td>0.28</td>
<td>0.44</td>
</tr>
</tbody>
</table>

Table 12: ergodic distribution normalized for every GDP class
C.2 Estimate with 3-year average growth rates

In this section we consider transitions with 3-year average growth rates, i.e. from $(y_t, \bar{g}_{t+2})$ to $(y_{t+3}, \bar{g}_{t+3,t+5})$, where $\bar{g}_{t+2}$ is the average annual growth rate from period $t$ to period $t+2$ and $\bar{g}_{t+3,t+5}$ is the average annual growth rate from period $t+3$ to period $t+5$. In the following we report the usual tables.

Table 13: distribution of the first and last year vs ergodic distribution for only GDP classes

<table>
<thead>
<tr>
<th>Year</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>1950</td>
<td>0.31</td>
<td>0.53</td>
<td>0.13</td>
<td>0.02</td>
</tr>
<tr>
<td>1998</td>
<td>0.16</td>
<td>0.38</td>
<td>0.22</td>
<td>0.25</td>
</tr>
<tr>
<td>Ergodic</td>
<td>0.05</td>
<td>0.17</td>
<td>0.22</td>
<td>0.57</td>
</tr>
</tbody>
</table>

Table 14: transition matrix

<table>
<thead>
<tr>
<th>States</th>
<th>I-</th>
<th>I+</th>
<th>I++</th>
<th>II-</th>
<th>II+</th>
<th>II++</th>
<th>III-</th>
<th>III+</th>
<th>III++</th>
<th>IV-</th>
<th>IV+</th>
<th>IV++</th>
</tr>
</thead>
<tbody>
<tr>
<td>120</td>
<td>0.50</td>
<td>0.28</td>
<td>0.23</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>137</td>
<td>0.23</td>
<td>0.52</td>
<td>0.19</td>
<td>0.01</td>
<td>0.02</td>
<td>0.02</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>121</td>
<td>0.12</td>
<td>0.17</td>
<td>0.42</td>
<td>0.11</td>
<td>0.10</td>
<td>0.07</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>268</td>
<td>0.04</td>
<td>0.02</td>
<td>0.02</td>
<td>0.49</td>
<td>0.23</td>
<td>0.19</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>255</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.29</td>
<td>0.42</td>
<td>0.28</td>
<td>0</td>
<td>0</td>
<td>0.01</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>340</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.14</td>
<td>0.19</td>
<td>0.55</td>
<td>0.04</td>
<td>0.01</td>
<td>0.07</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>83</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.04</td>
<td>0.05</td>
<td>0.05</td>
<td>0.35</td>
<td>0.18</td>
<td>0.34</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>73</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.21</td>
<td>0.27</td>
<td>0.47</td>
<td>0.04</td>
<td>0.01</td>
<td>0</td>
</tr>
<tr>
<td>197</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.13</td>
<td>0.17</td>
<td>0.56</td>
<td>0.03</td>
<td>0.04</td>
<td>0.09</td>
</tr>
<tr>
<td>55</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.07</td>
<td>0.02</td>
<td>0.02</td>
<td>0.36</td>
<td>0.29</td>
<td>0.24</td>
</tr>
<tr>
<td>93</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.20</td>
<td>0.47</td>
<td>0.32</td>
</tr>
<tr>
<td>88</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.16</td>
<td>0.41</td>
<td>0.43</td>
</tr>
</tbody>
</table>

Table 15: ergodic distribution

<table>
<thead>
<tr>
<th>I-</th>
<th>I+</th>
<th>I++</th>
<th>II-</th>
<th>II+</th>
<th>II++</th>
<th>III-</th>
<th>III+</th>
<th>III++</th>
<th>IV-</th>
<th>IV+</th>
<th>IV++</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>0.01</td>
<td>0.04</td>
<td>0.05</td>
<td>0.05</td>
<td>0.04</td>
<td>0.04</td>
<td>0.10</td>
<td>0.15</td>
<td>0.27</td>
<td>0.23</td>
<td></td>
</tr>
</tbody>
</table>
Table 16: ergodic distribution normalized for every GDP class

<table>
<thead>
<tr>
<th></th>
<th>-</th>
<th>+</th>
<th>++</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>0.35</td>
<td>0.36</td>
<td>0.30</td>
</tr>
<tr>
<td>II</td>
<td>0.32</td>
<td>0.30</td>
<td>0.38</td>
</tr>
<tr>
<td>III</td>
<td>0.27</td>
<td>0.21</td>
<td>0.51</td>
</tr>
<tr>
<td>IV</td>
<td>0.23</td>
<td>0.42</td>
<td>0.36</td>
</tr>
</tbody>
</table>

Table 17: distribution of the first and last year vs ergodic distribution for only GDP classes

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>1960</td>
<td>0.31</td>
<td>0.53</td>
<td>0.13</td>
<td>0.02</td>
</tr>
<tr>
<td>1998</td>
<td>0.16</td>
<td>0.39</td>
<td>0.20</td>
<td>0.25</td>
</tr>
<tr>
<td>Ergodic distr.</td>
<td>0.03</td>
<td>0.13</td>
<td>0.19</td>
<td>0.65</td>
</tr>
</tbody>
</table>

D Analytical model

This appendix presents the analytical model depicted in Figure 2. Consider an economy with the following typical Solovian capital accumulation equation:

\[ \dot{k} = sf(k) - (\delta + n)k, \]  

where \( k \) is capital per capita, \( s \) is the constant saving rate, \( f \) is the production function, \( \delta \) is the depreciation rate of capital, \( n \) is the growth rate of population. Under the following assumptions:

- \( f(0) = 0 \);
- \( f' > 0 \ \forall k > 0 \), \( \lim_{k \to 0} f' > \frac{n + \delta}{s} \) and \( \lim_{k \to +\infty} f' = a > \frac{n + \delta}{s} \);
- \( f'' > 0 \ \forall k \in [\hat{k}, \bar{k}] \) and \( f'' < 0 \ \forall k \in [0, \hat{k}] \sim [\bar{k}, +\infty) \)

the most interesting cases are two.

In the first case every country, independent of its initial (positive) level of capital, has a long-run growth rate of capital per capita equal to \( sa - n - \delta \). This happens if \( f(k)/k > (n + \delta)/s \ \forall k \in [0, \infty) \). The proof is straightforward from Equation (1), since \( \dot{k} = 0 \) for \( k = 0 \) and \( \dot{k} > 0 \) for \( k > 0 \). The change in concavity is the cause of the nonlinear pattern of the growth rate (which depends on average capital productivity).
In the second case there are two equilibria: \( \bar{k}_1 < \bar{k}_2 \). This happens if \( \exists \bar{k} \in (0, \infty) \) such that \( f(k)/k < (n + \delta)/s \). The first equilibrium \( \bar{k}_1 \) is an attractor: in fact, \( \dot{k} > 0 \) for \( k \in [0, \bar{k}_1) \) and \( \dot{k} < 0 \) for \( k \in (\bar{k}_1, \bar{k}_2) \) (this directly derives from the shape of \( f \)).

The second equilibrium, \( \bar{k}_2 \), is unstable: in fact, \( \dot{k} < 0 \) for \( k \in (\bar{k}_1, \bar{k}_2) \) and \( \dot{k} > 0 \) for \( k \in (\bar{k}_2, \infty) \). This implies that a country will converge to the equilibrium with a lower level of capital if its initial level of capital is lower than \( \bar{k}_2 \), while it will have a positive long-run growth equal to \( sa - n - \delta \) if its initial level of capital is greater than \( \bar{k}_2 \). Also in this case the growth rate can follow a nonlinear path.

Finally, restating the above results in term of per capita income we have:

\[
\frac{\dot{y}}{y} = \frac{f'(k)}{f(k)} \frac{\dot{k}}{k} = \frac{f'(k)}{f(k)} \left[ s \frac{f(k)}{k} - n - \delta \right].
\]

Thus, also \( \dot{y}/y \) has a non monotonic path and \( \lim_{k \to +\infty} \dot{k}/k = \dot{y}/y = sa - n - \delta \) since \( \lim_{k \to +\infty} \frac{f'(k)}{f(k)} = 1 \) for the assumption on \( f \) for \( k \to +\infty \). Figure 2 reports the relationship between the growth rate and the level of income: the growth path represented by a solid line and Trajectory A refer to the first case, while Trajectory B refers to the second case.

### E Inference on Markov transition matrices

In this appendix we illustrate a procedure to make inference on the elements of a Markov transition matrix.

#### E.1 Basic notation

Suppose that the observations of a process with \( k \) states, i.e. with state space \( S = 1, \ldots, k \), are collected for more than one period. Let \( n_{ij} \) be the number of observations in the sample corresponding to transitions from state \( i \) to state \( j \), \( n_i = \sum_{j=1}^{k} n_{ij} \) the total number of observations in state \( i \), and \( n_i = (n_{i1}, \ldots, n_{ik}) \) the vector collecting all \( n_{ij} \), \( i \in S \) and \( j = 1, \ldots, k \); hence \( n = \sum_{i=1}^{k} n_i \) is the total number of observations.

Let \( P \) be the \((k \times k)\) transition matrix. The element \( p_{ij} \) represents the transition probability from state \( i \) to state \( j \), so that \( \sum_{j=1}^{k} p_{ij} = 1 \) and \( 0 \leq p_{ij} \leq 1 \). Moreover, let \( p_i \) be the fraction of observations in initial state \( i \), i.e. \( p_i = n_i/n \).

Suppose the ergodic distribution for this process exists. The ergodic distribution is defined as:

\[
\pi = \pi P
\]

(2)
under the constraint:

\[ \pi u' = 1, \]

where \( u \) is the sum vector. From another point of view \( \pi \) corresponds to a row of the matrix \( P^t \) for \( t \to \infty \).

### E.2 Inference

In the following we assume that the rows of \( P \) are independent.

#### E.2.1 Consistent estimators

The maximum likelihood (ML) estimator of \( P, \hat{P}, \) is given by:

\[
\hat{P} = [\hat{p}_{ij}] = \left[ \frac{n_{ij}}{n_i} \right],
\]

where \( n_i = \sum_{j=1}^{n} n_{ij} \) (for a proof see e.g. Anderson and Goodman (1957)). \( \hat{P} \) being the ML estimator, these estimates are consistent.

In general, take \( P \) and a function \( M \) such that \( M : P \to \mathbb{R} \). Since \( P \) is unknown, then \( M (P) \) is unknown as well. A natural estimator is \( \hat{M} = M (\hat{P}) \), which, in turn, is consistent (see Trede (1999)). \( M \) can represent any function (linear and non-linear), e.g. the function which associates the transition matrix to an element of its ergodic distribution (when it exists).

#### E.2.2 Distribution of estimates

Stuart and Ord (1994), p. 260, show that the distribution of \( n_i \) converges to a \( n \)-variate normal distribution, with means \( n_i p_{ij} \), variances \( n_i p_{ij} (1 - p_{ij}) \) and covariances \( \text{cov} (n_{ij}, n_{iq}) = -n_i p_{ij} p_{iq} \). Thus \( \sqrt{n_i} (\hat{p}_{ij} - p_{ij}) \) tends towards the normal distribution \( N(0; p_{ij} (1 - p_{ij})) \). Notice that, defining \( p_i (k, \bar{k}) = \sum_{k=\bar{k}}^{\bar{k}} p_{ik} \), then \( \sqrt{n_i} (\hat{p}_i (k, \bar{k}) - p_i (k, \bar{k})) \) tends towards the normal distribution \( N(0; p_i (k, \bar{k}) (1 - p_i (k, \bar{k}))) \).

The asymptotic distribution of \( \hat{M} \) can be derived by the delta method (DM) (see Trede (1999)). Consider the first order Taylor series expansion of \( M (\hat{P}) \) around \( M (P) \):

\[
M (\hat{P}) = M (P) + DM (P) \left( \text{vec} \left( \hat{P}' - P' \right) \right),
\]

where

\[
DM (P) = \frac{\partial M (P)}{\partial \text{vec} (P')}
\]
E.2 Inference

is a $1 \times k^2$ vector, which contains the first derivatives of $M$ with respect to each element of $P$.

Since the rows of $P$ are independent and each row tends towards a $n$-variate normal distribution, we have

\[ \sqrt{n} \left( \text{vec} \left( \hat{P}' - P' \right) \right) \xrightarrow{d} N \left( 0, V \right) , \]

where

\[ V = \begin{bmatrix} V_1 & \cdots & \vdots \\ \vdots & \ddots & \vdots \\ \vdots & \cdots & V_k \end{bmatrix} \]

is block diagonal with

\[ V_m = [v_{m,ij}] = \begin{cases} \frac{p_{mi}(1-p_{mi})}{p_m} & \text{for } i = j \\ -\frac{p_{mi}p_{mj}}{p_m} & \text{for } i \neq j \end{cases} \]

for $m = 1, \ldots, k$ and 0 elsewhere.

Therefore the asymptotic distribution of $M$ is given by:

\[ \sqrt{n} \left( M \left( \hat{P} \right) - M \left( P \right) \right) \xrightarrow{d} N \left( 0, \sigma^2_M \right) , \]

where

\[ \sigma^2_M = \left( DM \left( P \right) \right) V \left( DM \left( P \right) \right)' . \]

Since both $DM \left( P \right)$ and $V$ are unknown, they are estimated by $DM \left( \hat{P} \right)$ and $\hat{V}$ calculated on the basis of (the elements of) $\hat{P}$. As $\hat{P}$ is a ML-estimator, then $DM \left( \hat{P} \right)$ and $\hat{V}$ are consistent too and therefore the estimate of the variance of $M$ is given by:

\[ \hat{\sigma}^2_M = \left( DM \left( \hat{P} \right) \right) \hat{V} \left( DM \left( \hat{P} \right) \right)' . \]

Since $M \left( P \right)$ is normally distributed, then the $(1 - \alpha)$-confidence interval for $M \left( \hat{P} \right)$ is

\[ M \left( \hat{P} \right) \pm c \frac{\hat{\sigma}_M}{\sqrt{n}} , \]

where $c$ is the $(1 - \frac{\alpha}{2})$-quantile of the $N \left( 0, 1 \right)$. Alternatively,

\[ s = \frac{M \left( \hat{P} \right) - M \left( P \right)}{\frac{\hat{\sigma}^2_M}{\sqrt{n}}} \]

converges towards a Gaussian distribution under the null hypothesis $M \left( \hat{P} \right) = M \left( P \right)$.
E.2.3 Testing

The Delta Method provides the most general procedure of testing. However, for the simpler tests on the elements of $P$ we use a more direct way: we focus on the comparison of two elements of the transition matrix and of two elements of ergodic distribution.

Tests on elements of $P$

Comparison of two elements of different rows The first test regards the difference between two transition probabilities belonging to different rows. Under the assumption of independence among the rows of $P$, $s = \frac{\hat{p}_{ij} - \hat{p}_{mq}}{\sqrt{\hat{\sigma}_{ij}^2/n_i + \hat{\sigma}_{mq}^2/n_m}}$ converges to a Gaussian distribution under the null hypothesis $\hat{p}_{ij} = \hat{p}_{mq}$, where $i \neq m$ and $\hat{\sigma}_{ij}^2 = \hat{p}_{ij}(1 - \hat{p}_{ij})$. The proof is straightforward given the normality of the asymptotic distribution of $P$ and the assumption of independence among the rows of $P$.

Comparison of two elements of the same row A second test regards the difference between two transition probabilities belonging to the same row. Then $s = \frac{\hat{p}_{ij} - \hat{p}_{iq}}{\sqrt{\hat{\sigma}_{ij}^2/n_i + \hat{\sigma}_{iq}^2/n_i - 2\text{cov}(\hat{p}_{ij}, \hat{p}_{iq})/n_i}}$ converges to a Gaussian distribution under the null hypothesis of an identical value of $\hat{p}_{ij}$ and $\hat{p}_{iq}$, where $j, q \in \{1, ..., n\}$ and $\text{cov}(\hat{p}_{ij}, \hat{p}_{iq}) = -\hat{p}_{ij}\hat{p}_{iq}$. Also in this case the proof is straightforward.

Starting from these two types of tests we can test all possible combinations among elements of $P$.

Tests on elements of $\pi$

Comparison between single elements of the ergodic distribution To test the difference between elements of the ergodic distribution requires the application of the DM. First we have to calculate the derivatives of the function $M$. In this case $M$ is a function calculating the difference between any two elements of this distribution. Let $ED(P)$ be this function:

$$ED(P) = \pi_q - \pi_m, \text{ where } m, q \in \{1, ..., k\}.$$ 

To calculate $\sigma_{ED}^2$ we need to know the analytical derivatives of the ergodic distribution with respect to the elements of the transition matrix. Conlisk (1985) provides an analytical formulation. Assume that the increase in the element $j$ in row $i$, $p_{ij}$, is
absorbed by a decrease in the element of the last column $k$ of row $i$, $p_{ik}$ (the row sum must sum to one). Thus, the derivative of the $q-th$ element of the ergodic distribution is defined as follows:

$$
\frac{\partial \pi_q}{\partial p_{ij}} = \pi_i \left( z_{jq} - z_{kq} \right) \forall i, j, q \in \{1, ..., k\},
$$

where $z_{jq}$ is an element of the fundamental matrix $Z = (I - P - bu')^{-1}$ and $b$ is any $1 \times k$ row vector such that $b'u \neq 0$.

Then:

$$
\frac{\partial ED (P)}{\partial p_{ij}} = \pi_i \left( z_{jq} - z_{kq} \right) - \pi_i \left( z_{jm} - z_{km} \right) = \pi_i \left[ (z_{jq} - z_{kq}) - (z_{jm} - z_{km}) \right],
$$

from which we can calculate $\partial^2 ED$. Applying (9) we obtain the confidence interval for $ED = \pi_q - \pi_m$ and/or by (10) we can test the null hypothesis $\pi_q = \pi_m$.

**Comparison between elements of the ergodic distribution normalized with respect to different subsets of states** Consider the following function of the transition matrix:

$$
EDN (P) = \frac{\pi_{q_1}}{\sum_{m=s_1}^{e_1} \pi_m} - \frac{\pi_{q_2}}{\sum_{m=s_2}^{e_2} \pi_m},
$$

where $q_1 \in \{s_1, ..., e_1\}$ and $q_2 \in \{s_2, ..., e_2\}$. This represents the difference in the elements of the ergodic distribution, normalized within two subsets of states; in our case these two subsets of states are $(s_1, ..., e_1)$ and $(s_2, ..., e_2)$.

The first step consists in calculating the first derivative of $EDN (P)$ with respect to the elements of $P$. Given the analytical derivative of an element of the ergodic distribution with respect to an element of $P$, when the last column $k$ absorbs any positive perturbation, we obtain:

$$
\frac{\partial EDN (P)}{\partial p_{ij}} = \frac{\partial \pi_{q_1}}{\partial p_{ij}} \left( \sum_{m=s_1}^{e_1} \pi_m \right) - \frac{\pi_{q_1}}{\left( \sum_{m=s_1}^{e_1} \pi_m \right)^2} \sum_{m=s_1}^{e_1} \frac{\partial \pi_m}{\partial p_{ij}} - \frac{\partial \pi_{q_2}}{\partial p_{ij}} \left( \sum_{m=s_2}^{e_2} \pi_m \right) + \frac{\pi_{q_2}}{\left( \sum_{m=s_2}^{e_2} \pi_m \right)^2} \sum_{m=s_2}^{e_2} \frac{\partial \pi_m}{\partial p_{ij}}
$$

Then, by (8) and (9) we can construct the confidence interval for $EDN (\hat{P})$ and/or by (10) we can test the null hypothesis $\pi_{q_1} / \left( \sum_{m=s_1}^{e_1} \pi_m \right) = \pi_{q_2} / \left( \sum_{m=s_2}^{e_2} \pi_m \right)$. 

36