Macroeconomic Dynamics, **20**, 2016, 1550–1580. Printed in the United States of America.
doi:10.1017/S1365100514001023

A TIME-VARYING MARKOV-SWITCHING MODEL FOR ECONOMIC GROWTH

Bruno Morier and Vladimir Kühl Teles
São Paulo School of Economics—FGV

This paper investigates patterns of variation in economic growth across and within countries using a time-varying transition matrix Markov-switching approach. The model developed here explains the dynamics of growth based on a collection of different states that countries pass into and out of over time; in addition, these states are characterized by their own submodels and growth patterns. The transition matrix among the different states varies over time—depending on the conditioning variables of each country—with a linear dynamic for each state. We develop a generalization of Diebold's EM algorithm and estimate a sample model in a panel with a transition matrix conditioned on institutional quality and the investment level. We find three states of growth: stable growth, miraculous growth, and stagnation. The results show that institutional quality is an important determinant of long-term growth, whereas the investment level plays a variety of roles: it contributes positively in countries with high-quality institutions but is of little relevance in countries with medium- or low-quality institutions.

Keywords: Growth Regimes, Markov Switching

1. INTRODUCTION

Economic growth varies across countries and over time. Several articles have attempted to explain the differences in cross-country growth rates and have shown that countries tend to converge on different growth regimes [Durlauf and Johnson (1995); Desdoigts (1999); Masanjala and Papageorgiou (2004); Tan (2010)]. Other articles have shown that a country can move among different regimes of growth, alternating between "miracle" and "failure" regimes [Easterly et al. (1993); Pritchett (2000, 2003); Hausmann et al. (2004); Jerzmanowski (2006); Jones and Olken (2008); Berg et al. (2012); Cuberes and Jerzmanowski (2012)]. This paper proposes a methodology for studying the differences in countries' growth rates and changes in their growth processes over time using a Markov regime-switching approach. However, rather than assuming—as is typical—that some countries are simply fortunate enough to have a "good transition probability matrix" and that others are

Address correspondence to: Vladimir K. Teles, São Paulo School of Economics—FGV, Rua Itapeva 474, São Paulo, SP 01332-000, Brazil; e-mail: vladimir.teles@fgv.br.

© 2015 Cambridge University Press 1365-1005/15

https://doi.org/10.1017/S1365100514001023 Published online by Cambridge University Press

born with a "bad transition matrix," we develop a generalization of Diebold's EM algorithm and estimate a time-varying transition matrix Markov-switching model. In this way, we estimate the parameters of the transition matrix as a function of time, in which the matrix is conditioned on a country's institutional quality and investment level.

The analyses performed in most studies do not incorporate all the available information related to growth and do not typically consider the intertemporal factor. More recently, however, several studies have focused on precisely this aspect by dating regime changes and finding explanatory variables. Jones and Olken (2008) used the methodology developed by Bai and Perron (2003) to date breaks in the growth process and included a descriptive panel with means and standard deviations for changes in variables correlated with growth, which was then incorporated into a Solow residual model that described the anatomy of the variations in growth. Hausmann et al. (2004) analyzed the transition to a better growth pattern (the acceleration event), selected one ad hoc criterion for acceleration, and performed various calculations using means and standard deviations for periods prior to and following the acceleration event. Similarly, Wacziarg and Welch (2003) investigated the average impact on economic growth of removing restrictions on external trade. Additional studies using the same approach include Aizenman and Spiegel (2007) and Jong-A-Pin and de Haan (2007). However, although this approach explores the intertemporal aspect, the data are not treated with a panel; in general, descriptive analyses are performed to determine the average impact of the variables under consideration.

According to Durlauf and Johnson (1995), traditional linear regressions demand a high degree of homogeneity in the production function among different countries, where this is always identical and in the Cobb–Douglas format. However, different countries have different growth patterns, and their production functions may be quite heterogeneous, which suggests the need for a more flexible approach. Durlauf and Johnson (1995) rejected a linear and homogeneous Cobb–Douglas function in favor of a model with multiple growth regimes. In contrast to the approach in our study, the regimes in Durlauf and Johnson (1995) were fixed for each country according to its initial conditions from the beginning to the end of the sample. This approach generated a series of studies on convergence clubs and growth regimes.

Durlauf and Johnson (1995) employed a regression tree technique to identify these regimes, whereas other studies have suggested alternative techniques, using fixed regimes for each country. Hansen (2000) developed a statistical theory of threshold estimation in which the asymptotic distribution is known, thus allowing testing. Masanjala and Papageorgiou (2004) used this theory to perform these tests, finding nonlinearity in the production function and a constant elasticity substitution (CES) format, which might explain the heterogeneity of the coefficients and the existence of multiple regimes. Durlauf and Johnson (1995) found four regimes. Desdoigts (1999) and Kourtellos (2002) used the projection pursuit technique to approximate the growth functions by projections onto lower dimension functions. Bloom et al. (2003), Paap et al. (2005), Alfo et al. (2008), and Basturk et al.

https://doi.org/10.1017/S1365100514001023 Published online by Cambridge University Press

(2008) estimate the growth regime that characterizes a country, along with other parameters, using a mixture analysis with classical statistics; Ardic (2006) uses the Bayesian statistic, whereas Canova (2004) uses the predictive density technique. In addition, Tan (2010) revisits the regression tree technique using generalized unbiased interaction detection and estimation (GUIDE).[1]

The fixed regime approach, however, does not explain the inconsistencies in growth or the breaks in regimes that characterize the experiences of developing countries and that strongly contrast with the experiences of the most advanced industrialized countries, upon which classical theories of economic growth are based. Pritchett (2000) examines this inconsistency among different countries and identifies six different growth patterns to analyze the dynamics of the growth level, along with breaks and variations. Pritchett (2003) suggests using an approach that features six different states of growth through which countries pass, where each state is associated with a level of growth determined by a different submodel, and the growth process is explained by toy collection models. This approach focuses on these changes in states—which are less common in developed countries—to explain the inconsistencies and breaks that developing countries experience. With this approach, multiple equilibria are permitted, including poverty traps, heterogeneous production functions, and different roles for the conditioning variables, which are nonlinear and have breaks. These factors are supported by strong empirical evidence. The changes in states are modeled probabilistically, with the probabilities conditioned on the prevailing foundational elements of each country. In an experiment with a constant ad hoc transition matrix, Pritchett (2003) managed to recreate the dispersion of the current per capita GDP levels.

The Pritchett (2003) approach was used in Jerzmanowski (2006), in which a Markov-switching model was estimated using a constant measure of the quality of the institutions as an explanatory variable for the transition matrix.[2] Jerzmanowski (2006) found a positive relationship between this variable and growth. Each state was modeled as a constant AR(1). Kerekes (2009) used a Markov-switching model for a similar classification of convergence clubs, and considered a small number of transition matrices that were constant and unconditioned; furthermore, each country was associated with a matrix. Both approaches follow the proposal of Pritchett (2003), although with considerably less model complexity than found in the original proposal and with fewer explanatory variables, which is perhaps due to the difficulties in modeling and in estimation, in addition to a lack of data.

This work proposes a Markov-switching model with a transition matrix that is conditional on the fundamental characteristics of each country and that varies over time. In each state, the resulting submodel is modeled as an AR(p) with exogenous variables. The possibility that the transition matrix will vary over time and that each submodel will possess a linear specification with exogenous variables allows a description that is closer to the approach of Pritchett (2003), which is the main contribution of this study. The model permits the following:

https://doi.org/10.1017/S1365100514001023 Published online by Cambridge University Press

1. Ruptures in the growth process;
2. Explanatory variables that vary among countries and over time;
3. Variables capable of affecting growth in various ways that are asymmetric and nonlinear (i.e., depending on the current state or that influence the transition among states); and
4. Multiple equilibria and clubs.

Based on this model, an estimation method was developed that generalized the EM algorithm described in Diebold et al. (1993). The original model permitted two states with a normal, unconditional distribution. In contrast, in the model developed here, there are no restrictions on the number of states, and each state is an AR(p) with exogenous variables.

Using this method, we performed estimations using a sample model in which the transition matrix is conditioned on institutional quality and the investment level. The primary results indicate that institutional quality is an important determinant of long-term growth, whereas the level of investment plays a differentiated role that contributes positively to growth in countries with high-quality institutions but is of little relevance in countries with medium- or low-quality institutions.

This study is organized as follows. The next section presents the model. Section 3 presents the estimation algorithm. Section 4 presents the empirical model and discusses the results. Finally, Section 5 concludes this study and provides a discussion for future studies.

2. THE MODEL

In this section, we present the general model with regime changes for growth. Consider a generic country, c. Let g_{ct} be the growth rate for country c between times $t-1$ and t. g_{ct} is modeled using a time-varying Markov-switching process with S AR(p) states. The precise meaning of this is that for each time t, country c is in an (unobserved) state $s_{ct} \in \{0, \cdots, S-1\}$, s_{ct} follows a Markov process with S states, and in each state, g_{ct} follows a different AR(p) process.

The exogenous variables may affect the growth process through the distribution of each AR(p) state or based on a transition matrix of the Markov process of s_{ct}. Let $z^{ct} = (z_i^{ct})_{i=1\ldots N_z}$ be the vector of the N_z exogenous variables that affect the growth process in each AR(p) state with $z_{1ct} = 1$. The same variables influence the growth of each country, although the value of these variables varies with c and t. For each state, the AR(p) process is given by the following expression:

$$g_{ct} = \left(\sum_{i=1}^{p} \alpha_{is_{ct}} L^i\right) \cdot g_{ct} + \left(\sum_{i=1}^{N_z} \gamma_{is_{ct}} \cdot z_i^{ct}\right) + \epsilon_{ct} \quad \text{with} \quad \epsilon_{ct} \sim N(0, \sigma_{s_{ct}}^2) \text{ (i.i.d.).} \tag{1}$$

https://doi.org/10.1017/S1365100514001023 Published online by Cambridge University Press

The coefficients $(\alpha, \gamma) = (\{\alpha_{is_{ct}}\}, \{\gamma_{is_{ct}}\})$ are dependent on the given state. α determines the degree to which the current growth is affected by the previous growth, thus constituting the only memory in the model beyond the previous state s_{t-1}. γ represents the degree to which each variable z_i influences the growth in each state. Finally, $\sigma = (\{\sigma_{s_{ct}}\})$ determines the standard deviation of each state.

The dynamics of s_{ct} are given by a Markov process with a transition matrix that varies over time and that is conditional on exogenous explanatory variables. These N_x conditional variables are collected in the vector $x^{ct} = (x_i^{ct})_{i=1\ldots N_x}$, with $x_{1ct} = 1$. We can write

$$P(s_{ct} = i | s_{ct-1} = j) = p_{ij}(x^{ct}). \tag{2}$$

Following Diebold et al. (1993) and Jerzmanowski (2006), we use a logit specification for $p_{ij}(x^{ct})$:

$$p_{ij}(x^{ct}) = \begin{cases} \dfrac{\exp(\beta_{ij} x^{ct})}{1 + \sum_{k=0}^{S-2} \exp(\beta_{kj} x^{ct})}, & \text{if } i = 0, \cdots, S-2 \\ 1 - \sum_{k=0}^{S-2} p_{kj}(x^{ct}), & \text{if } i = S-1. \end{cases} \tag{3}$$

The coefficients $\beta = \{\beta_{ij}\}$ determine how each exogenous variable in x_k influences the probability of transition between states i and j. Each β_{ij} is an $(N_x \times 1)$ vector.

This model is flexible, with the explanatory variables conditioning the process of each state and the transition matrix. This model is more realistic than the usual Markov-switching specification with a fixed matrix [as in Hamilton (1994)] because it allows the fundamental characteristics of each country to influence the dynamics of the state. It is reasonable to assume, for example, that the probability of a crisis in a country with high-quality institutions and good macroeconomic policies is lower than that in a country with low-quality institutions and poor policies. The fundamental characteristics may change over time, as is the case with investments and political regimes.

We now develop the immediate results of the model. Letting $P_{ct} = P_{ct}(x^{ct})$ be a matrix with coefficients $p_{ij}(x^{ct})$ and letting $\hat{s_{ct}} = \left[P(s_{ct} = 0|\hat{\theta}), \ldots, P(s_{ct} = S-1|\hat{\theta})\right]'$ be the probability of occurrence for each state for some information set $\hat{\theta}$, we obtain the following result:

$$\hat{s_{ct}} = P_{ct} \hat{s}_{ct-1}, \tag{4}$$

https://doi.org/10.1017/S1365100514001023 Published online by Cambridge University Press

given by the Markov property of s_{ct}. Long-term growth is well-defined for each state if z_i are stationary and if the coefficients do not imply a unit root. In this case, letting $\mu_{z_i} = E[z_i]$ be the expectation value for each exogenous variable, the long-term growth for each state s_t is given by

$$\text{long-run growth of } s_{ct} : E[g_{ct}|s_{ct} = j] = \frac{\sum_{i=1}^{N_z} (\gamma_{ij}\mu_{z_i})}{\left(1 - \sum_{i=1}^{p} \alpha_{ij}\right)}. \tag{5}$$

Maintaining the transition matrix P_{ct} as a constant (which can be interpreted as maintaining the same policies and the status quo), we can use equation (4) to derive the unconditional probabilities for each state. Taking $\pi = (\pi_0, \cdots, \pi_{S-1})'$ as the unconditional probability, equation (4) implies that $\pi = P_{ct}\pi$. We also have $\sum_i \pi_i = 1$. As in Hamilton (1994), the solution to this system is given by

$$\pi = (A'A)^{-1}Ae_{S+1}, \tag{6}$$

where e_{S+1} denotes the $S + 1$ column of the identity matrix I_{S+1} and where

$$\underset{(S+1 \times S)}{A} = \begin{bmatrix} I_S - P_{ct} \\ 1 \cdots 1 \end{bmatrix}. \tag{7}$$

In this case, we can use equations (5) and (6) to calculate the long-term growth of the economy of country c with

$$E[g_{ct}] = \sum_{i=0}^{S-1} E[g_{ct}|s_t = i]\pi_i. \tag{8}$$

3. ESTIMATION METHOD

In this section, we extend the algorithm from Diebold et al. (1993) for the model described in the preceding section. The algorithm is based on the expectation maximization technique, which is flexible and appropriate for new extensions. A discussion on the EM algorithm can be found in Dempster et al. (1977), Watson (1983), and Ruud (1991). The following studies use Markov-switching algorithms that use the EM method: Hamilton (1990), Durland and McCurdy (1994), and Filardo (1994).

Following the model described in the preceding section, the data available for the estimation should be g_{ct}, z^{ct}, and x^{ct}. We assume that these data are available

https://doi.org/10.1017/S1365100514001023 Published online by Cambridge University Press

between $t = 1$ and $t = T$. From here on, we will use underlined notation to indicate the collection of all the data available for one variable up until the time index, such that $\underline{x_c}_T = (x_{c1}, \ldots, x^{ct})$. Thus, the input for the estimation is $\underline{g_c}_T, \underline{x_c}_T, \underline{z_c}_T$.

Let $\rho_j^c = P(s_{c1} = j)$, $\rho^c = (\rho_0^c, \ldots, \rho_{S-1}^c)$ and $\rho = (\{\rho^c\})$. The state of each country, s_{ct}, is hidden, and the probability of occurrence of each state follows the dynamics described in (4). Thus, the joint distribution of s_{ct} is determined by ρ_c and by the transition matrix P_{ct}.

The parameters for the estimation are collected in $\theta = (\alpha, \beta, \gamma, \sigma, \rho)$. These parameters will be the final output of the estimation algorithm. If states s_{ct} were known, we could simply estimate the complete data function of maximum likelihood to obtain estimates for the parameters of the model. Even though this is not the case, the calculation of this function proves useful. The function will be calculated by iterating a dynamic model following equation (4) and using the conditional distribution for each state given in (1). Because the growth of each country in the model is independent, we will first set a country c and calculate its function of maximum likelihood. For lighter notation, we will remove c and use g_t rather than g_{ct}. We then have

$$
\begin{aligned}
&f(\underline{g}_T, \underline{s}_T | \underline{x}_T, \underline{z}_T, \theta) \\
&\quad = f(g_1, s_1 | \underline{x}_T, \underline{z}_T, \theta) \prod_{t=2}^{T} f(g_t, s_t | \underline{g}_{t-1}, \underline{s}_{t-1}, \underline{x}_T, \underline{z}_T, \theta) \\
&\quad = f(g_1 | s_1, \underline{x}_T, \underline{z}_T, \theta) P(s_1) \prod_{t=2}^{T} f(g_t | \underline{s}_t, \underline{g}_{t-1}, \underline{x}_T, \underline{z}_T, \theta) \\
&\qquad \times P(s_t | \underline{s}_{t-1}, \underline{g}_{t-1}, \underline{x}_{t-1}, \underline{z}_{t-1}, \theta) \\
&\quad = f(g_1 | s_1, z_1, \theta) P(s_1) \prod_{t=2}^{T} f(g_t | s_t, \underline{g}_{t-1}, z_t, \theta) P(s_t | s_{t-1}, x_t, \theta), \qquad (9)
\end{aligned}
$$

where f is the probability density given by (1) such that

$$
\begin{aligned}
&f(g_t | s_t = j, \underline{g}_{t-1}, z_t, \theta) \\
&\quad = \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp \left\{ \frac{-\left[g_t - \left(\sum_{i=1}^{p} \alpha_{is_t} L^i \right) \cdot g_t - \left(\sum_{i=1}^{N_z} \gamma_{is_t} \cdot z_{it} \right) \right]^2}{2\sigma_j^2} \right\}. \qquad (10)
\end{aligned}
$$

We can use equations (4), (9), and (10) to calculate the complete-data function of maximum likelihood for a country. We can further express this function of maximum likelihood in terms of an indicator function, separating the possible

https://doi.org/10.1017/S1365100514001023 Published online by Cambridge University Press

realizations of the states for each time t. Thus, we obtain

$$f(\underline{g}_T, \underline{s}_T|\underline{x}_T, \underline{z}_T, \theta)$$
$$= \sum_{j=0}^{S-1} f(g_1|s_1 = j, z_1, \theta)\rho_j I(s_1 = j) \prod_{t=2}^{T} \sum_{j_1=0}^{S-1} \sum_{j_2=0}^{S-1} I(s_t = j_1, s_{t-1} = j_2)$$
$$\times f(g_t|s_t = j_1, \underline{g}_{t-1}, z_t, \theta) P(s_t = j_1|s_{t-1} = j_2, x_t, \theta), \quad \textbf{(11)}$$

where $\rho_j = P(s_1 = j)$ is the jth coordinate of ρ and the indicator function I is 1 if the condition is true and 0 if it is not. Then, applying the logarithm and the expectation operator $E[.]$ in equation (11), we obtain

$$E\left\{\log\left[f(\underline{g}_T, \underline{s}_T|\underline{x}_T, \underline{z}_T, \theta)\right]\right\} = \sum_{j=0}^{S-1} \rho_j \left[\log f(g_1|s_1 = j, z_1, \theta) + \log \rho_j\right]$$
$$+ \sum_{t=2}^{T} \sum_{j_1=0}^{S-1} \sum_{j_2=0}^{S-1} P(s_t = j_1, s_{t-1} = j_2|\underline{g}_T, \underline{x}_T, \underline{z}_T, \theta)$$
$$\times \left[\log f(g_t|s_t = j_1, \underline{g}_{t-1}, z_t, \theta) + \log P(s_t = j_1|s_{t-1} = j_2, x_t, \theta)\right]. \quad \textbf{(12)}$$

We observe that the operator $E[.]$ essentially substitutes $I(.)$ for $P(.)$ in the equation. Using the independence assumption, the log maximum likelihood function for the panel of countries is obtained by the sum of the log maximum likelihood for each country. Let Ψ be this function, and let ψ_c be the log maximum likelihood function for each country. We then have

$$\Psi = \sum_c \Psi_c = \sum_c \sum_{j=0}^{S-1} \rho_j \left[\log f(g_{c1}|s_{c1} = j, z_{c1}, \theta) + \log \rho_{cj}\right]$$
$$+ \sum_c \sum_{t=2}^{T} \sum_{j_1=0}^{S-1} \sum_{j_2=0}^{S-1} P(s_{ct} = j_1, s_{ct-1} = j_2|\underline{g_c}_T, \underline{x_c}_T, \underline{z_c}_T, \theta)$$
$$\times \left[\log f(g_{ct}|s_{ct} = j_1, \underline{g_c}_{t-1}, z^{ct}, \theta) + \log P(s_{ct} = j_1|s_{ct-1} = j_2, x^{ct}, \theta)\right]. \quad \textbf{(13)}$$

The last equation is the basis of the estimation procedure. We did not compute the complete maximum likelihood function because we did not observe the states that were realized. However, this last equation for the expectation of the complete function of maximum likelihood can be estimated. The parameters in θ will be estimated by maximising equation (13).

https://doi.org/10.1017/S1365100514001023 Published online by Cambridge University Press

3.1. EM Algorithm: General Procedure

Given the model from Section 2, we have created formulae for all the terms of equation (13), with the exception of $P(s_{ct} = j_1, s_{ct-1} = j_2|\underline{g_c}_T, \underline{x_c}_T, \underline{z_c}_T, \theta)$. This last term refers to the smoothed probability of the occurrence of states $s_{ct} = j_1, s_{ct-1} = j_2$, which is conditional on all the data previous and subsequent to time t. We can obtain this term algorithmically. Because the term is conditional on all the data, it tends to experience variations that are smaller than those of the other terms of the equation for values of θ that are near the maximum of the expected maximum likelihood.

The concept of optimization thus emerges in two stages, in which the equation (13) is maximized by maintaining the smoothed probabilities as constant:

(1) Select a $\theta^{(0)}$
(2) With the last $\theta^{(l)}$ available, calculate $P(s_t = j_1, s_{t-1} = j_2|\underline{g}_T, \underline{x}_T, \underline{z}_T, \theta^{(l)})$ for $(j_1, j_2) \in (0, \ldots, S-1)^2$ and $t \in (2, \ldots, T)$
(3) Calculate $\theta^{(l+1)} = \arg\max_{\theta} E\left\{\log\left[f_l(\underline{g}_T, \underline{s}_T|\underline{x}_T, \underline{z}_T, \theta)\right]\right\}$
(4) If $||\theta^{(l+1)} - \theta^{(l)}|| > \epsilon$, then return to step (2),

where ϵ and the $||.||$ norm define a convergence criterion. f_l from stage (3) is obtained via the smoothed probabilities calculated with the last available $\theta^{(l)}$, i.e., by replacing $P(s_{ct} = j_1, s_{ct-1} = j_2|\underline{g_c}_T, \underline{x_c}_T, \underline{z_c}_T, \theta)$ with $P(s_{ct} = j_1, s_{ct-1} = j_2|\underline{g_c}_T, \underline{x_c}_T, \underline{z_c}_T, \theta^{(l)})$. Stage (2) is called the expectation step, and stage (3) is called the maximization step. When there is convergence, the θ found by the procedure will be a numerical approximation to the maximum of equation (13).

3.2. Expectation Step

The algorithm described here calculates the smoothed probabilities of the occurrence of $s_{ct} = j_1, s_{ct-1} = j_2$ for all pairs (j_1, j_2) with an optimization described in Diebold et al. (1993). Again, we will treat each country separately, thus removing c from the notation:

1. Precalculus
   (a) Use equation (10) to calculate $f(g_t|s_t, \underline{g}_{t-1}, \underline{x}_t, \underline{z}_t, \theta^{(l)}) = f(g_t|s_t, \underline{g}_{t-1}, z_t, \theta^{(l)})$ for $(t, s_t) \in (1, \ldots, T) \times (0, \ldots, S-1)$
   (b) Use equation (3) to calculate the matrix P_{kt}
2. Calculate the filtered probabilities[3] $P(s_t, s_{t-1}|\underline{g}_t, \underline{x}_t, \underline{z}_t, \theta^{(l)})$ for $(s_t, s_{t-1}) \in (0, \ldots, S-1)^2$ and $t \in (2, \ldots, T)$
   (a) Calculate $f(g_t, s_t, s_{t-1}|\underline{g}_{t-1}, \underline{x}_t, \underline{z}_t, \theta^{(l)})$: For $t = 2$, we use the identity

$$f(g_2, s_2, s_1|\underline{g}_1, \underline{x}_2, \underline{z}_2, \theta^{(l)}) = \underbrace{f(g_2|s_2, \underline{g}_1, \underline{x}_2, \underline{z}_2, \theta^{(l)})}_{\text{Step 1}} \underbrace{P(s_2|s_1, \underline{g}_1, \underline{x}_2, \underline{z}_2, \theta^{(l)})}_{\text{Step 1}} \underbrace{P(s_1)}_{\rho_{s_1}}. \quad (14)$$

For $t = 3, \ldots, T$, we use the identity

https://doi.org/10.1017/S1365100514001023 Published online by Cambridge University Press

$$f(g_t, s_t, s_{t-1}|\underline{g}_{t-1}, \underline{x}_t, \underline{z}_t, \theta^{(l)}) = \sum_{s_{t-2}=0}^{S-1} \underbrace{f(g_t|s_t, \underline{g}_{t-1}, \underline{x}_t, \underline{z}_t, \theta^{(l)})}_{\text{Step 1}} \underbrace{P(s_t|s_{t-1}, \underline{g}_{t-1}, \underline{x}_t, \underline{z}_t, \theta^{(l)})}_{\text{Step 1}} \underbrace{P(s_{t-1}, s_{t-2}|\underline{g}_{t-1}, \underline{x}_t, \underline{z}_t, \theta^{(l)})}_{\text{Step 2.a above}}. \tag{15}$$

(b) Calculate the probability density of g_t:

$$f(g_t|\underline{g}_{t-1}, \underline{x}_t, \underline{z}_t, \theta^{(l)}) = \sum_{s_t=0}^{S-1} \sum_{s_{t-1}=0}^{S-1} f(g_t, s_t, s_{t-1}|\underline{g}_{t-1}, \underline{x}_t, \underline{z}_t, \theta^{(l)}). \tag{16}$$

(c) The filtered probabilities are obtained for time t:

$$P(s_t, s_{t-1}|\underline{g}_t, \underline{x}_t, \underline{z}_t, \theta^{(l)}) = \frac{f(g_t, s_t, s_{t-1}|\underline{g}_{t-1}, \underline{x}_t, \underline{z}_t, \theta^{(l)})}{f(g_t|\underline{g}_{t-1}, \underline{x}_t, \underline{z}_t, \theta^{(l)})}. \tag{17}$$

3. Calculate the smoothed probabilities $P(s_t, s_{t-1}|\underline{g}_T, \underline{x}_T, \underline{z}_T, \theta^{(l)})$ for $(s_t, s_{t-1}) \in (0, \ldots, S-1)^2$ and $t \in (2, \ldots, T)$

(a) For each time $t \in (2, \ldots, T)$ and each pair $(s_t, s_{t-1}) \in (0, \ldots, S-1)^2$, sequentially calculate $P(s_\tau, s_{\tau-1}, s_t, s_{t-1}|\underline{g}_{t-1}, \underline{x}_{t-1}, \underline{z}_{t-1}, \theta^{(l)})$ for all $\tau \in (t+2, \ldots, T)$. For $\tau = t+1$, we use the identity

$$P(s_{t+1}, s_t, s_{t-1}|\underline{g}_{t+1}, \underline{x}_{t+1}, \underline{z}_{t+1}, \theta^{(l)}) = \frac{f(g_{t+1}|s_{t+1}, \underline{g}_t, \underline{x}_{t+1}, \underline{z}_{t+1}, \theta^{(l)}) P(s_{t+1}|s_t, \underline{g}_t, \underline{x}_{t+1}, \underline{z}_{t+1}, \theta^{(l)}) P(s_t, s_{t-1}|\underline{g}_t, \underline{x}_{t+1}, \underline{z}_{t+1}, \theta^{(l)})}{f(g_{t+1}|\underline{g}_t, \underline{x}_{t+1}, \underline{z}_{t+1}, \theta^{(l)})}, \tag{18}$$

where the first two terms of the numerator were obtained from the first step, the third term was obtained from step 2(c), and the denominator was obtained from step 2(b)

For $\tau \in (t+2, \ldots, T)$, we use the identity

$$P(s_\tau, s_{\tau-1}, s_t, s_{t-1}|\underline{g}_\tau, \underline{x}_\tau, \underline{z}_\tau, \theta^{(l)}) = \sum_{s_{\tau-2}=0}^{S-1} f(g_\tau|s_\tau, \underline{g}_{\tau-1}, \underline{x}_\tau, \underline{z}_\tau, \theta^{(l)}) \times \frac{P(s_\tau|s_{\tau-1}, \underline{g}_{\tau-1}, \underline{x}_\tau, \underline{z}_{\tau-1}, \theta^{(l)}) P(s_{\tau-1}, s_{\tau-2}, s_t, s_{t-1}|\underline{g}_{\tau-1}, \underline{x}_\tau, \underline{z}_\tau, \theta^{(l)})}{f(g_\tau|\underline{g}_{\tau-1}, \underline{x}_\tau, \underline{z}_\tau, \theta^{(l)})}, \tag{19}$$

where the first two terms of the numerator were obtained in the first step, the third term of the numerator was obtained from step 3(a), and the denominator was obtained from step 2(b)

(b) Calculate the smoothed probabilities for time t:

$$P(s_t, s_{t-1}|\underline{g}_T, \underline{x}_T, \underline{z}_T, \theta^{(l)}) = \sum_{s_T=0}^{S-1} \sum_{s_{T-1}=0}^{S-1} P(s_T, s_{T-1}, s_t, s_{t-1}|\underline{g}_T, \underline{x}_T, \underline{z}_T, \theta^{(l)}). \tag{20}$$

https://doi.org/10.1017/S1365100514001023 Published online by Cambridge University Press

3.3. Maximization Step

To maximize equation (13), we will compute the first-order conditions in relation to $(\alpha, \beta, \gamma, \sigma)$. As suggested in Hamilton (1994) for a model with a fixed P_{ct}, we calculate ρ from equation (6), i.e., as the long-term expectation for the probability of occurrence of each state given the conditions of the exogenous variables and the current estimation of β. Because P_{ct} varies in this model, the option is given by the matrix P_{c1}.

Let

$$\begin{aligned}
\Psi^l &= E\left\{ \log\left[f_l(\underline{g}_T, \underline{s}_T | \underline{x}_T, \underline{z}_T, \theta) \right] \right\}, \\
\varphi_j^{ct} &= \log f(g_{ct} | s_{ct} = j, \underline{g_c}_{t-1}, z^{ct}, \theta), \\
\xi_{j_1,j_2}^{ct} &= P(s_{ct} = j_1, s_{ct-1} = j_2 | \underline{g_c}_T, \underline{x_c}_T, \underline{z_c}_T, \theta^{(l)}), \\
\xi_j^{ct} &= P(s_{ct} = j | \underline{g_c}_T, \underline{x_c}_T, \underline{z_c}_T, \theta^{(l)}), \\
p_{j_1,j_2}^{ct} &= P(s_{ct} = j_1 | s_{ct-1} = j_2, x^{ct}, \theta), \\
\Gamma_j^{ct} &= g_{ct} - \left(\sum_{i=1}^{p} \alpha_{i,j} L^i \right) \cdot g_{ct} - \left(\sum_{i=1}^{N_z} \gamma_{i,j} \cdot z_i^{ct} \right).
\end{aligned} \tag{21}$$

We observe that the ξs are constant in relation to $(\alpha, \beta, \gamma, \sigma)$, which are estimates. This relationship is given by

$$\xi_j^{ct} = \begin{cases} \sum_{j_2=0}^{S-1} \xi_{j_1,j_2}^{ct}, \text{ if } t > 1 \\ \rho_j^c, \text{ if } t = 1. \end{cases} \tag{22}$$

Applying the log in equation (10), we obtain

$$\varphi_j^{ct} = \frac{-\left[\Gamma_j^{ct}\right]^2}{2\sigma_j^2} - \frac{\log\left(2\pi\sigma_j^2\right)}{2}. \tag{23}$$

Using equation (22) and the preceding definitions (21), we obtain a simpler representation for Ψ^l that is derived from equation (13) and from replacing the dependent smoothed probabilities on θ with $\theta^{(l)}$:

$$\begin{aligned}
\Psi^l &= \sum_c \sum_{j=0}^{S-1} \rho_j \left(\varphi_j^{ct} + \log \rho_j^c\right) + \sum_c \sum_{t=2}^{T} \sum_{j_1=0}^{S-1} \sum_{j_2=0}^{S-1} \xi_{j_1,j_2}^{ct} \left(\varphi_j^{ct} + \log p_{j_1,j_2}^{ct}\right) \\
&= \sum_{c,t,j} \xi_j^{ct} \varphi_j^{ct} + \sum_{c,j} \rho_j^c \log \rho_j^c + \sum_{t=2}^{T} \sum_{c,j_1,j_2} \xi_{j_1,j_2}^{ct} \log p_{j_1,j_2}^{ct}.
\end{aligned} \tag{24}$$

https://doi.org/10.1017/S1365100514001023 Published online by Cambridge University Press

In the second line of this equation, we have a representation in which the first term is a function of (α, γ, σ), the second term is a function of ρ, and the third term is a function of β:

1. Calculation of α and of γ:
   By manipulating equation (23), we have

$$\frac{\partial \varphi_{j_1}^{ct}}{\partial \alpha_{k,j}} = \begin{cases} \dfrac{L^k g_{ct} \Gamma_j^{ct}}{\sigma_j^2}, & \text{if } j_1 = j \\ 0, & \text{if } j_1 \neq j, \end{cases} \tag{25}$$

$$\frac{\partial \varphi_{j_1}^{ct}}{\partial \gamma_{k,j}} = \begin{cases} \dfrac{z_k^{ct} \Gamma_j^{ct}}{\sigma_j^2}, & \text{if } j_1 = j \\ 0, & \text{if } j_1 \neq j. \end{cases} \tag{26}$$

   By manipulating equation (24) and substituting equations (25) and (26), we obtain the first-order conditions

$$\frac{\partial \Psi^l}{\partial \alpha_{k,j}} = \frac{\sum_{c,t} \xi_j^{ct} L^k g_{ct} \left[g_{ct} - \left(\sum_{i=1}^{p} \alpha_{i,j} L^i \right) \cdot g_{ct} - \left(\sum_{i=1}^{N_z} \gamma_{i,j} \cdot z_i^{ct} \right) \right]}{\sigma_j^2} = 0, \tag{27}$$

$$\frac{\partial \Psi^l}{\partial \gamma_{k,j}} = \frac{\sum_{c,t} \xi_j^{ct} z_k^{ct} \left[g_{ct} - \left(\sum_{i=1}^{p} \alpha_{i,j} L^i \right) \cdot g_{ct} - \left(\sum_{i=1}^{N_z} \gamma_{i,j} \cdot z_i^{ct} \right) \right]}{\sigma_j^2} = 0. \tag{28}$$

   Rearranging the last equation, we obtain

$$\sum_{i=1}^{p} \left[\sum_{c,t} \xi_j^{ct} (L^k g_{ct})(L^i g_{ct}) \right] \cdot \alpha_{i,j} + \sum_{i=1}^{N_z} \left[\sum_{c,t} \xi_j^{ct} (L^k g_{ct}) z_i^{ct} \right] \cdot \gamma_{i,j} = \sum_{c,t} \xi_j^{ct} (L^k g_{ct}) g_t, \tag{29}$$

$$\sum_{i=1}^{p} \left[\sum_{c,t} \xi_j^{ct} z_k^{ct} (L^i g_{ct}) \right] \cdot \alpha_{i,j} + \sum_{i=1}^{N_z} \left[\sum_{c,t} \xi_j^{ct} z_k^{ct} z_i^{ct} \right] \cdot \gamma_{i,j} = \sum_{c,t} \xi_j^{ct} z_k^{ct} g_{ct}. \tag{30}$$

   For each fixed j, we have a linear system with the variables $(\alpha_{i,j}, \gamma_{i,j})$, with $p + N_z$ variables and with $p + N_z$ equations (varying the k in the two equations). $(\alpha_{i,j}, \gamma_{i,j})$ are calculated by solving this system.
2. Calculation of σ:
   By manipulating equation (23), we obtain

$$\frac{\partial \varphi_{j_1}^{ct}}{\partial \sigma_j} = \begin{cases} \dfrac{\left(\Gamma_j^{ct}\right)^2}{\sigma_j^3} - \dfrac{1}{\sigma_j}, & \text{if } j_1 = j \\ 0, & \text{if } j_1 \neq j. \end{cases} \tag{31}$$

   By manipulating equation (24) and substituting equations (31), we obtain the first-order conditions

$$\frac{\partial \Psi^l}{\partial \sigma_j} = \sum_{c,t} \xi_j^{ct} \cdot \left(\frac{\left(\Gamma_j^{ct}\right)^2}{\sigma_j^3} - \frac{1}{\sigma_j} \right) = 0. \tag{32}$$

https://doi.org/10.1017/S1365100514001023 Published online by Cambridge University Press

The volatilities of each state may then be calculated with

$$\sigma_j = \frac{\sum_{c,t} \xi_j^{ct} \left(\Gamma_j^{ct}\right)^2}{\sum_{c,t} \xi_j^{ct}}. \tag{33}$$

3. Calculation of β:
Manipulating equation (24), we obtain

$$\frac{\partial \Psi}{\partial \beta_{i_1,i_2}} = \sum_{t=2}^{T} \sum_{c,j_2} \frac{\partial}{\partial \beta_{i_1,i_2}} \sum_{j_1=0}^{S-1} \xi_{j_1,j_2}^{ct} \log p_{j_1,j_2}^{ct}. \tag{34}$$

Note that

$$\begin{aligned} 1 - \sum_{k=0}^{S-2} p_{k,j_2}^{ct} &= \left[1 + \sum_{k=0}^{S-2} \exp(\beta_{k,j_2} x^{ct})\right]^{-1}, \\ \frac{\partial}{\partial \beta_{i_1,i_2}} \log \left[1 + \sum_{k=0}^{S-2} \exp(\beta_{k,i_2} x^{ct})\right] &= x^{ct} p_{i_1,i_2}^{ct}. \end{aligned} \tag{35}$$

We can expand the term that is derived using equation (3). Using the first identity of equation (35),

$$\begin{aligned} &\sum_{j_1=0}^{S-1} \xi_{j_1,j_2}^{ct} \log p_{j_1,j_2}^{ct} \\ &= \sum_{j_1=0}^{S-2} \xi_{j_1,j_2}^{ct} \log \left[\frac{\exp(\beta_{j_1,j_2} x^{ct})}{1 + \sum_{k=0}^{S-1} \exp(\beta_{k,j_2} x^{ct})}\right] + \xi_{S-1,j_2}^{ct} \log \left(1 - \sum_{k=0}^{S-2} p_{k,j_2}^{ct}\right) \\ &= \sum_{j_1=0}^{S-2} \xi_{j_1,j_2}^{ct} \beta_{j_1,j_2} x^{ct} - \xi_{j_2}^{ct} \log \left[1 + \sum_{k=0}^{S-2} \exp(\beta_{k,j_2} x^{ct})\right]. \end{aligned} \tag{36}$$

Substituting this into equation (34) and using the second identity of equation (35), we obtain the first-order condition

$$\frac{\partial \Psi^l}{\partial \beta_{i_1,i_2}} = \sum_{c} \sum_{t=2}^{T} x^{ct} \left(\xi_{i_1,i_2}^{ct} - \xi_{i_2}^{ct-1} p_{i_1,i_2}^{ct}\right) = 0, \tag{37}$$

with which it is possible to determine β. However, equation (37) is not linear because of the option arising from probabilities expressed in logit form. We can, however, approximate p_{i_1,i_2}^{ct} in a first-order Taylor series from $\beta^{(l)}$ in a manner that is analogous to that of Diebold et al. (1993). This approximation will be particularly good as we approach the convergence criterion, i.e., the real β:

$$p_{i_1,i_2}^{ct} = p_{i_1,i_2}^{ct}(\beta_{i_1,i_2}) = p_{i_1,i_2}^{ct}(\beta^{(l)}) + \phi_{i1,i2}^{ct} \cdot (\beta_{i_1,i_2} - \beta_{i_1,i_2}^{(l)})', \tag{38}$$

https://doi.org/10.1017/S1365100514001023 Published online by Cambridge University Press

where $p^{ct}_{i_1,i_2}(.)$ designates the probability of transition in the function of $p^{ct}_{i_1,i_2}$, as in equation (3). The derivative is given by

$$\phi^{ct}_{i1,i2} = \frac{\partial p^{ct}_{i_1,i_2}}{\partial \beta_{i_1,i_2}}(\beta^{(l)}) = \frac{\partial}{\partial \beta_{i_1,i_2}} \left[\frac{\exp(\beta_{i_1,i_2} x^{ct})}{1 + \sum_{k=0}^{S-1} \exp(\beta_{k,i_2} x^{ct})} \right]$$

$$= (x^{ct})' \left[p^{ct}_{i_1,i_2}(\beta^{(l)}) - p^{ct}_{i_1,i_2}(\beta^{(l)})^2 \right]. \quad (39)$$

Substituting equation (38) into equation (37) and rearranging, we obtain

$$\beta_{i_1,i_2} = \beta^{(l)}_{i_1,i_2} + \left(\sum_{c,t} x^{ct} \xi^{ct-1}_{i_2} \phi^{ct}_{i_1,i_2} \right)^{-1} \left\{ \sum_{c,t} x^{ct} \left[\xi^{ct}_{i_1,i_2} - \xi^{ct-1}_{i_2} p^{ct}_{i_1,i_2}(\beta^{(l)}) \right] \right\}, \quad (40)$$

with which we can calculate the beta.

4. EMPIRICAL MODEL

In this section, we use the general model and the estimation algorithm developed in the last two sections to estimate an exemplary empirical model with a transition matrix that varies over time and that is conditioned on institutional quality and by the investment level (%GDP). Each state is modeled following an AR(1) without explanatory exogenous variables.

The resulting model has three states, which is one less than is found in Jerzmanowski (2006). We are unaware of an objective statistical criterion in the literature for selecting the number of states in Markov-switching models similar to the one used here. However, the selection is performed because, with one variable and the conditioning from the transition probabilities, the estimated model with four states appears to require too much of the available data; in other words, its resulting transition matrix presents many cases of 100% or 0% transition probabilities, which indicates few occurrences of transitions that are conditional on the effective values of the pair of explanatory variables.

Despite having one less state, the results corroborate those of the Jerzmanowski (2006) model with virtually the same first three states from that model: stable growth, miraculous growth, and stagnation. The only differences appear to be that the stagnation state in this study is a mix of the crisis and stagnation states from Jerzmanowski (2006) because of the greater volatility that is presented. The resulting model can be seen as a test of the robustness of the Jerzmanowski (2006) findings because the results are analogous to the introduction of investment with a different data set: that author's data were from 89 countries during a period from 1962 to 1994 and were taken from Penn World Tables, version 6.1.

We can compare the states that were found with the six states that were established in an exogenous and ad hoc manner in Pritchett (2003). Stable growth

https://doi.org/10.1017/S1365100514001023 Published online by Cambridge University Press

appears to be the union of "steady growth at levels of income near world leaders" and "nonconverging steady growth.' The miraculous growth state corresponds to the rapid growth state. The stagnation state corresponds to the poverty trap, non-poverty-trap stagnation, and growth implosion states.

4.1. Data

Data from 73 countries over the period 1962–2007 were considered.[4]

To measure institutional quality, we used the index of government anti-diversion policies (GADP) from Hall and Jones (1999) and Jerzmanowski (2006). This index follows the same principle as the index found in Knack and Keefer (1995) and combines a measurement of the observance of laws, risk of expropriation, corruption, and the quality of bureaucracy, among others. The index is assumed to be constant over the time period and to vary only by country, as in Jerzmanowski (2006). This variable will be denoted as GADP_c.

To measure the level of investment, we use gross fixed capital formation (GFCF) from the World Bank database (WDI). For each time t, the conditioning variable of the transition matrix is the mean of the GFCF for the five previous years $(t-5,\ldots,t-1)$. The resulting five-year average is a slowly changing measure and is calculated using times that precede t to avoid endogeneity problems, which are typically associated with studies involving growth and investment. In this case, current investment only affects the expectancy for changes in future growth states, beyond current growth state model assessment—through the transition matrix. For some countries in the sample, an extrapolation for the initial years was performed because of a lack of data. However, the GFCF series exhibits slow changes during this time, and even considering that the series used is the average of five years, we should not be biased by the procedure adopted. Henceforth, this variable (average GFCF of the five previous lags) will be denoted by GFCF_{ct}. For economic growth, the per capita GDP growth from the Penn World Tables (PWT), version 6.3, was used.

4.2. Results

In this subsection, we discuss the results. We begin with the main results of the estimation (estimated parameters), and in sequence, we present a brief discussion with examples from specific countries.

Table 1 presents the parameters of the AR(1) process described by equation (1) for each state $j=(1,2,3)$. The first column shows the constant term (γ_{1j}), the second column shows the coefficient of lag 1 (α_{1j}), and the third column shows the volatility (σ_j). Long-term growth is calculated with equation (5).

State 1, which we call stable growth, is the state that best represents the experience of developed countries with a growth rate close to 2%, a high autocorrelation

https://doi.org/10.1017/S1365100514001023 Published online by Cambridge University Press

TABLE 1. States: AR(1) Process

State	Constant (γ_{1j})	AR[1] (α_{1j})	Volatility (σ_j)	Growth long term
1	1.27	0.35	2.42	1.97
2	4.20	0.30	3.46	6.03
3	0.15	0.04	6.47	0.16

among growth rates, and low volatility. Countries such as Colombia and Pakistan, with stable and nonconvergent growth as defined in Pritchett (2003), also fit into this category.

State 2, which we call miraculous growth, is a state that tends to have a short duration in most countries and consists mostly of "catch-ups" in growth by less developed countries. In the available historical data, the countries that maintained this state for the longest period of time are the emerging Asian countries, with Japan and China as illustrative cases.

State 3, which we call stagnation, is most common among emerging countries that fail to achieve convergence and among poor African countries. It is a highly persistent state for the poorest countries and is characterized by very high volatility; countries in this state achieve some years of strong growth, but they generally end up with years of negative growth, and average growth is quite low.

Tables 2–4 present data for the most important examples of countries and periods of incidence for each of the states. The analysis is descriptive, and we have qualified a variable as "High" or "Low" if the mean in each table has a standard score equal to or greater than 0.7 (high) or equal to or less than −0.7 (low).

The incidence in each state in these tables follows the convention of classifying countries and periods within each state when reaching a probability greater than 75% for occurrence for each of the states. The variables of GADP and GFCF are included, with values that are normalized to the mean and standard deviation for the entire sample, which includes all countries (standard score). The GFCF value of each country presented in the table is the mean of the GFCF normalized for the period in which the state in question occurs.

The threshold for including a country in each table is the number of years associated with the state being analyzed from summing all episodes. Because state 1 (Table 2) is more stable, a 30-year threshold was chosen (67% of the time sample). A total of 27 countries are included in this table (38% of the sample). For state 2 (Table 3), the less stable state, a 10-year threshold was chosen (22% of the sample). A total of 12 countries are included in this table (17% of the sample). For state 3 (Table 4), the median state in terms of stability, a 20-year threshold was chosen (44% of the sample). A total of 14 countries are included in this table (20% of the sample).

https://doi.org/10.1017/S1365100514001023 Published online by Cambridge University Press

TABLE 2. Main countries and periods of incidence of state 1

Country	GADP	GFCF	Periods		
Australia	1.29	0.74	1962–		
Austria	1.38	0.46	1962–1975	1981–	
Belgium	1.40	0.42	1962–		
Bolivia	−1.25	−0.73	1972–1980	1986–	
Canada	1.50	−0.01	1962–1981	1984–	
Switzerland	1.61	0.89	1962–1974	1978–	
Colombia	−0.40	−0.40	1963–		
Costa Rica	0.09	−0.30	1963–1979	1985–	
Denmark	1.54	0.03	1965–		
Spain	0.70	0.38	1974–		
Finland	1.52	0.50	1962–1967	1973–1990	1996–
France	1.34	0.12	1962–		
United Kingdom	1.30	−0.43	1962–		
Greece	0.28	0.78	1977–		
India	−0.28	−0.01	1972–1998	2000–	
Israel	0.49	0.23	1975–		
Italy	0.76	0.03	1963–1971	1980–	
Japan	1.30	1.09	1977–		
Sri Lanka	−0.87	0.04	1963–1971	1981–	
Netherlands	1.56	0.23	1962–		
Norway	1.46	0.79	1962–		
New Zealand	1.55	0.21	1962–1966	1978–	
Pakistan	−0.91	−0.58	1963–1968	1973–	
El Salvador	−1.29	−0.93	1963–1977	1983–	
Sweden	1.55	0.01	1962–		
United States	1.37	−0.37	1962–		
South Africa	0.41	0.01	1963–		
Mean	0.72	0.12			

Table 2 shows that the mean of the normalized GADP is high, i.e., 0.72, which indicates a relationship between this variable and stable growth. In addition, most countries that do not have a high GADP have episodes in which there is a deviation from this state. Some of these countries, such as Colombia, Pakistan, and India prior to 1999, are explicitly classified as "stable and nonconvergent growth" by Pritchett (2003). This status also appears to apply to other countries with a below-average GADP. The table does not reveal a strong relationship between the GFCF and the occurrence of state 1.

Table 3 shows that the mean for GFCF is high, i.e., 0.96, which indicates a relationship between this variable and miraculous growth. In addition, most of the countries that do not have a high GFCF maintain miraculous growth for a few years. Asian countries tend to maintain this type of growth. The level of investment decreased in Greece after the occurrence of this state. Japan's per capita

https://doi.org/10.1017/S1365100514001023 Published online by Cambridge University Press

TABLE 3. Main countries and periods of incidence of state 2

Country	GADP	GFCF	Periods		
Brazil	0.14	−0.08	1968–1978		
Chile	−0.02	−0.28	1977–1979	1986–1995	
China Version 1	−0.05	1.89	1971–		
Greece	0.28	3.40	1963–1973		
Hong Kong	0.65	0.66	1963–1988	1999–	
Indonesia	−0.77	−0.72	1969–1979		
Japan	1.30	1.96	1963–1973		
Korea, Republic of	0.39	0.75	1963–1979	1982–1996	
Malaysia	0.17	0.73	1973–1980	1987–1996	
Portugal	0.74	0.61	1963–1973	1976–1977	1986–1988
Singapore	0.96	1.75	1965–2003		
Thailand	0.28	0.87	1966–1995		
Mean	0.34	0.96			

TABLE 4. Main countries and periods of incidence of state 3

Country	GADP	GFCF	Periods			
Argentina	−0.33	0.05	1963–1966	1978–1991	1999–2002	
Central African Republic	−1.07	−1.11	1967–1997			
Cote d'Ivoire	−0.11	−0.46	1964–1968	1979–1993		
Fiji	−0.18	−0.21	1965–1968	1973–1975	1978–2000	
Gambia, The	−0.38	−0.91	1966–1967	1970–1972	1974–1983	1990–2001
Honduras	−1.05	−0.08	1969–1971	1978–1988	1990–1995	1998–2000
Iran	−0.76	0.56	1974–1995			
Mali	−1.57	−0.66	1963–1990	1994–1997		
Niger	−0.63	−1.38	1963–2004			
Senegal	−0.76	−1.37	1963–1985			
Togo	−0.95	−0.18	1970–2003			
Trinidad and Tobago	−0.16	0.23	1968–1994			
Venezuela	−0.18	0.47	1974–2004			
Congo, Dem. Rep.	−1.97	−1.48	1963–2005			
Mean	−0.72	−0.47				

GDP converged to that of the leading countries, thus concluding the "catch-up" phase, which also began stable growth after a transition period of stagnation. The GADP was also above average. The exception of the table appears to be Indonesia, which had a single miraculous growth event that lasted 10 years.

We observe in Table 4 that the mean GFCF is low: −0.72. We thus observe a relationship between these variables and the occurrence of stagnation. The GFCF is also quite low: −0.47.

https://doi.org/10.1017/S1365100514001023 Published online by Cambridge University Press

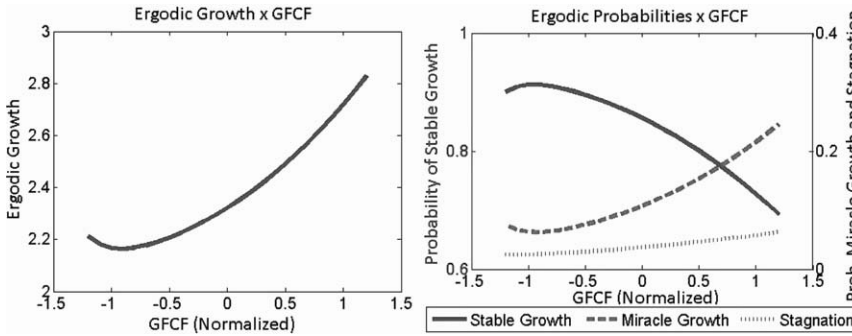


FIGURE 1. The Impact of investment for countries with high-quality institutions. Normalized GADP set at 1.2.

The role of institutional quality and investment level. In this subsection, we explain the relationships among institutional quality, investment, and the transition matrix. As discussed in Section 1, the model that is adopted allows a more complete description of the impact of each variable than a purely linear model.

Because two variables are under consideration, we first normalized the level of the institutions at three levels (−1.2,0,1.2), and we analyzed the impact of investment at each of these levels. The levels were selected for observation of the dynamics not only for countries with average institutions but also for those with high-quality or low-quality institutions. The selected levels correspond approximately to the 10th and 90th percentiles. Next, we established the same three levels of investment for a similar analysis.

Figure 1 shows the impact of investment for countries with high-quality institutions. The graph on the left shows the ergodic growth (long-term), and the graph on the right shows the ergodic probabilities (unconditional) of occurrence for each state. These are calculated from equations (5), (6), and (8). There is a positive impact on investment: the highest levels of investment result in a greater amount of time allocated to the miraculous growth state, generating greater long-term growth.

Figure 2 explains the changes in the transition matrix caused by variations in the levels of investment. Each graph indicates the probability of occurrence of each state in $t + 1$, conditional on the current state in t. The figure shows that the greatest impact of investment is found in the maintenance of miraculous growth (first graph), with no relevant role observed for the other situations. The probability of maintaining this growth increases from 53% for the countries with the least amount of investment to approximately 75% for the countries with the greatest amount of investment.

Figure 3 presents the case of medium-quality institutions. In this case, increased investment does not significantly impact growth. The effect even appears to be slightly negative because of the amount of time in stagnation. However, the effect appears to be somewhat irrelevant and generates a variation of 0.3%

https://doi.org/10.1017/S1365100514001023 Published online by Cambridge University Press

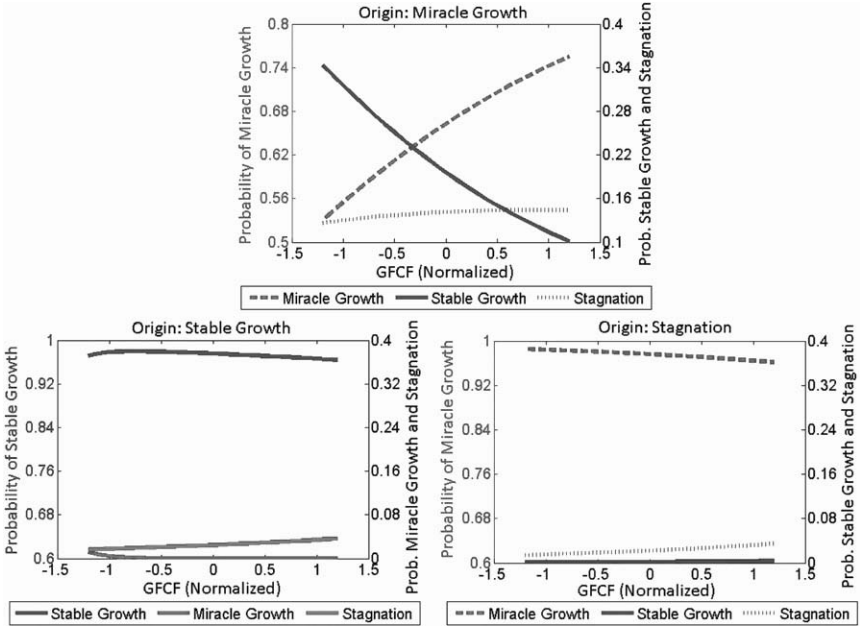


FIGURE 2. The impact of investment for countries with high-quality institutions. Normalized GADP set at 1.2.

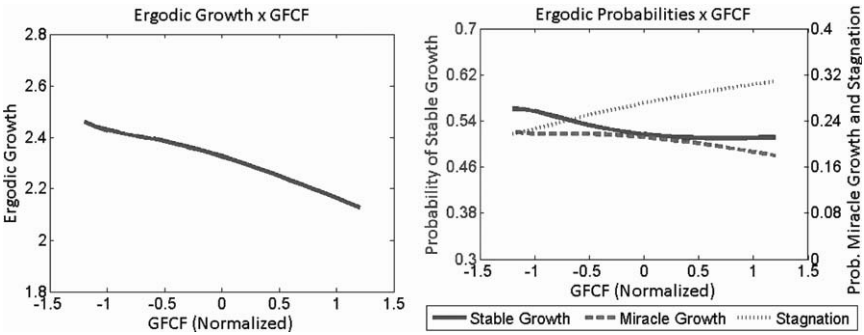


FIGURE 3. The impact of investment for countries with medium-quality institutions. Normalized GADP set at 0.

from end to end in the level of investment. At this level of institutional quality, miraculous growth is frequent but is counterbalanced in terms of average growth by stagnations, which are also frequent. The growth is more volatile than in the previous case, and a significant portion of the incidence of miraculous growth is due to the "catch-ups" from previous stagnations and not "catch-ups" in income level, as was predominant in the previous case.

https://doi.org/10.1017/S1365100514001023 Published online by Cambridge University Press

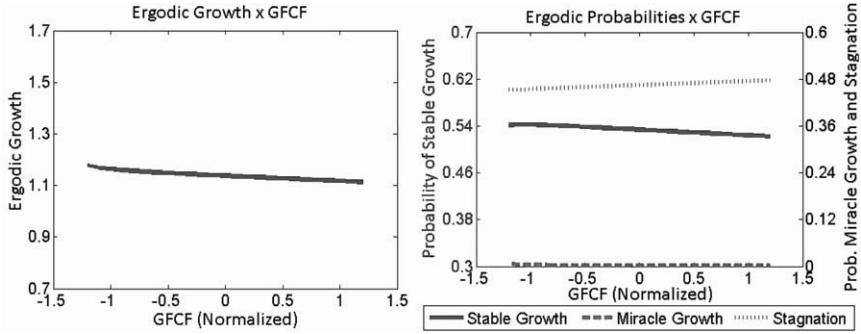


FIGURE 4. The impact of investment for countries with low-quality institutions. Normalized GADP set at -1.2.

Finally, Figure 4 presents the case of poor institutional quality. In this case, the growth dynamics is not favorable, and the level of investment does not have a positive effect. On the contrary, the effect is irrelevant, is less than in the previous case, and even tends to be negative.

This last finding is interesting. High levels of current investment do not seem to lead to better future growth prospects for countries with medium- and low-quality institutions, although investment can still have an effect on current growth (not measured by the model). The effect of above-average investment is unclear and can be slightly negative for average institutions.

The diverse effects of investment on growth in countries with different levels of institutional quality can be explained by the varied incentives to allocate investment in environments characterized by different institutional quality. Poor institutional quality may direct investment toward activities that do not always lead to increased productivity or that are not most appropriate to promote economic growth.

Thus, Ehrlich and Lui (1999) claim that corruption and inefficient bureaucracy tend to divert investment into unproductive purposes rather than into developing human capital. In addition, poor institutions may lead to misallocated investment that is directed by loss of some sectors, reducing the total productivity factor [Hsieh and Klenow (2009); Jones (2011)]. This argument is even stronger for public investment.

We continue with the analysis of a fixed level of investment. Figure 5 shows the dynamic with respect to a high level of investment in which ergodic growth is mostly increasing and institutions strongly influence growth. In general, higher quality institutions increase the likelihood of miraculous growth and decrease stagnation.

The exception is found for the final plot of the level of institutional quality of above 0.7, in which there is a decline in miraculous growth. This effect, which also occurs in the Jerzmanowski (2006) model, is given by the correlation between high institutional quality and the per capita GDP level. In the sample, there is a concentration of countries with high per capita GDP among the countries with

https://doi.org/10.1017/S1365100514001023 Published online by Cambridge University Press

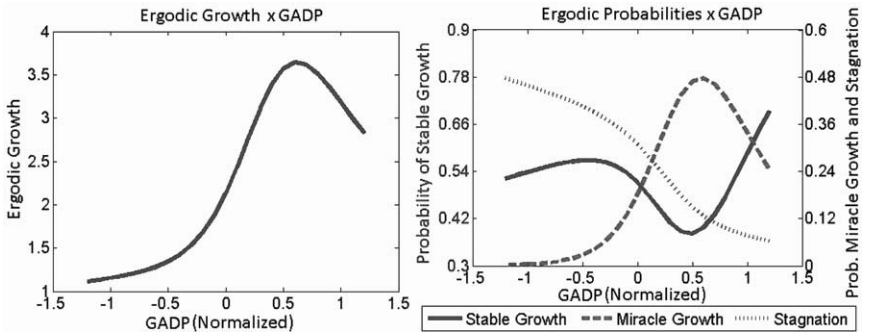


FIGURE 5. The impact of institutions for high-investment countries. Normalized GFCF set at 1.2.

very high quality institutions, and high per capita GDP seems to decrease the incidence of miraculous growth. According to Pritchett (2003), this result is due to the "catch-up" in growth convergence.

Following the reasoning of Pritchett (2003), a possible explanation for the correlation between the institutional level and per capita GDP in the sample is that the initial conditions of per capita GDP are given by accumulated growth until the beginning of the sample and with institutional quality as a variable that is related to growth. Accumulated growth is influenced by this variable. However, it is logical to include per capita GDP in the transition matrix, although perhaps only in the transition from the miraculous growth state.

Figure 6 presents a transition matrix with a high level of investment for the function of institutional quality. We can find the reasons for shorter stagnation periods in the first two tables. When experiencing stagnation, an economy with better institutional quality has an increased chance of recovery and moving into miraculous growth, whereas a country with low-quality institutions has a high probability of remaining stagnant, i.e., in a poverty trap. When in a state of steady growth, an economy with better institutions experiences greater stability, whereas an economy with lower institutional quality has a greater chance of falling into stagnation, although there might be a long period during which it maintains stable growth. The third table shows this effect, which has been discussed previously: miraculous growth continues for a shorter period of time and segues into stable growth. Thus, the increase in the probability of stagnation is minimal and immaterial because the incidence of stagnation for countries with the highest levels of institutional quality is not perennial; in addition, it is more likely that miraculous growth will occur, which will be even faster as the level of the institutions increases.

Figures 7 and 8 show the cases of average and below-average levels of investment. The dynamic is similar to that discussed in the case of above-average investment but with smaller oscillations in the incidence of miraculous growth and, therefore, smaller oscillations in ergodic growth.

https://doi.org/10.1017/S1365100514001023 Published online by Cambridge University Press

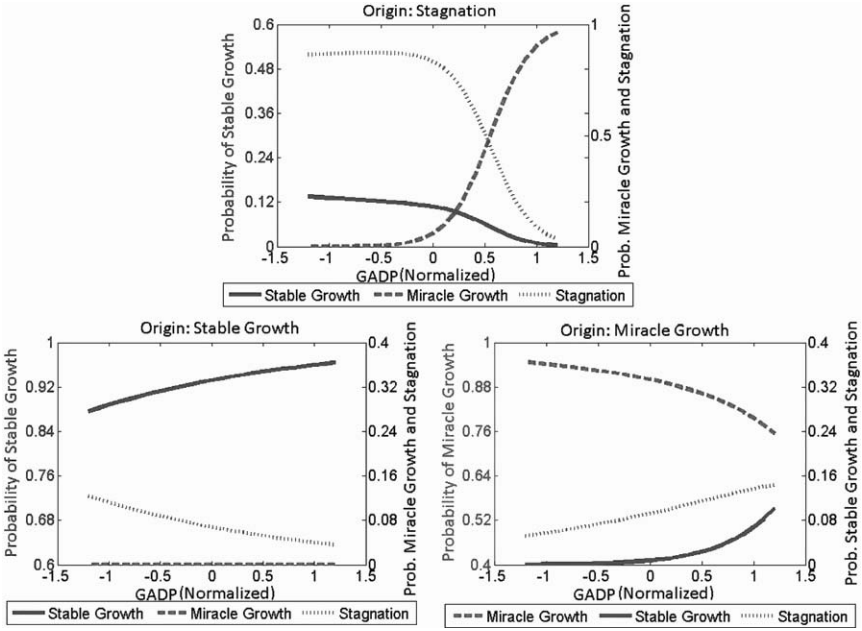


FIGURE 6. The impact of institutions for high-investment countries. Normalized GFCF set at 1.2.

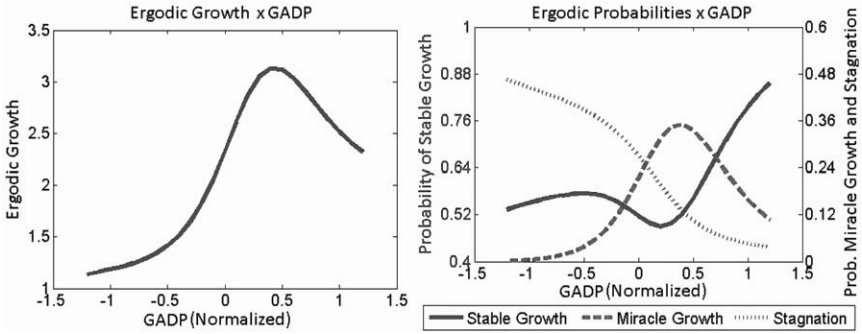


FIGURE 7. The impact of institutions for medium-investment countries. Normalized GFCF set at 0.

Cases of selected countries. In this section, we present the states generated by the model for six selected countries: Brazil, Argentina, China, Japan, the United States, and the Democratic Republic of the Congo. The objective is to use practical cases to demonstrate how regimes change and to show that certain factors may impact these changes. The model seeks to explain the probabilities of change that

https://doi.org/10.1017/S1365100514001023 Published online by Cambridge University Press

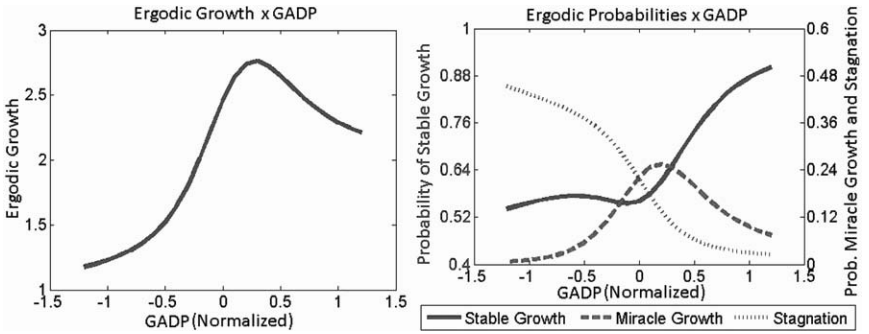


FIGURE 8. The impact of investment for low-investment countries. Normalized GFCF set at -1.2.

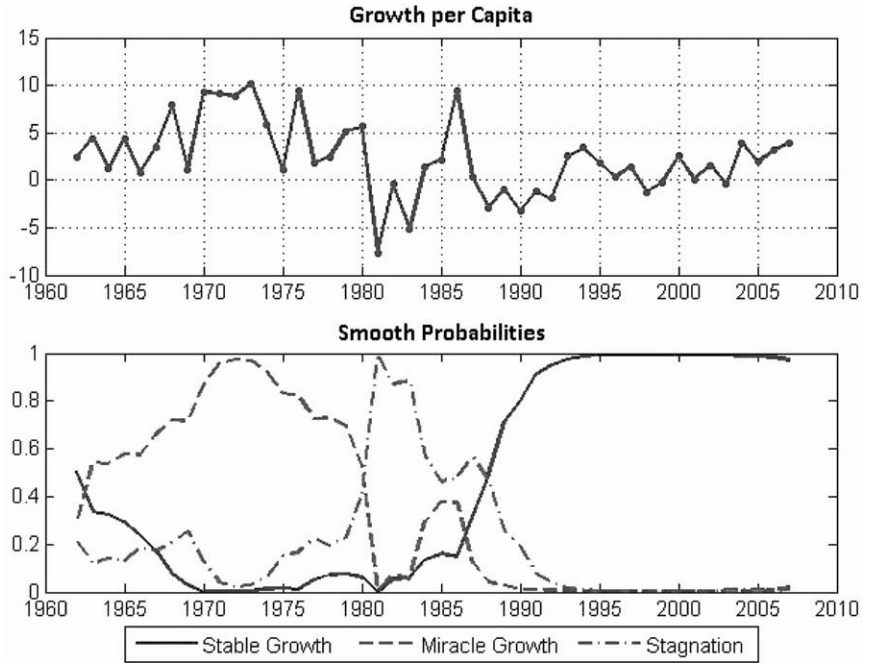


FIGURE 9. Growth and the occurrence of states for Brazil.

are conditioned on the structural variables, but most of the changes have cyclical and specific aspects that lead to the transitions.

In Figure 9, we examine Brazil. In the 1960s and the 1970s, we observe a dominant probability of miraculous growth. Despite rising debt and the

https://doi.org/10.1017/S1365100514001023 Published online by Cambridge University Press

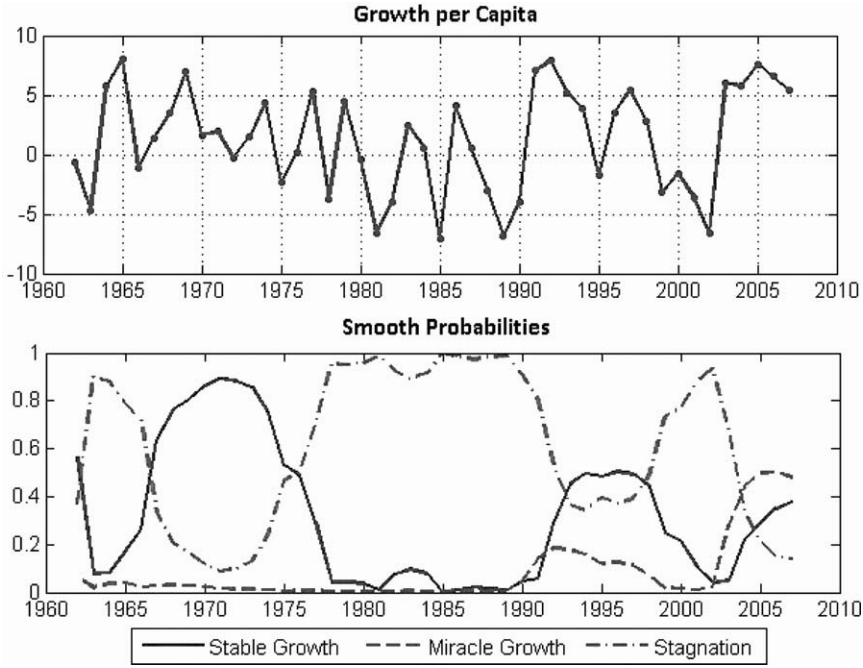


FIGURE 10. Growth and the occurrence of states for Argentina.

political and social turmoil that characterized this period, there was strong growth spurred by investment, particularly investment in infrastructure. In the 1980s, the effects of the oil crisis, which occurred before the 1980s, could no longer be circumvented. In a world that was overwhelmed by the oil crisis, Brazil had to implement adjustments for previous excesses in its external and fiscal accounts, while simultaneously controlling for inflation. In addition, this occurred at a time of political transition with the end of the dictatorship. After the constitution of 1988 and the implementation of inflationary controls, there was stabilization of growth at a nonconvergent stable state, to apply the terminology of Pritchett (2003).

Figure 10 shows the case of Argentina. According to Pritchett (2003), Argentina, for the most part, faces stagnation but cannot be classified as a poverty trap. Argentina follows the growth pattern of many Latin countries that experienced a prolonged period of stagnation after the oil crisis. After some degree of recuperation during the reforms of the Menem era, a new crisis period was experienced that featured external deficits, capital flight, and debt defaults.

In Figure 11, we present the impressive case of China. With the highest level of investment as a percentage of GDP, China has maintained miraculous growth since the 1970s. The most rapid growth process began with the reforms during the

https://doi.org/10.1017/S1365100514001023 Published online by Cambridge University Press

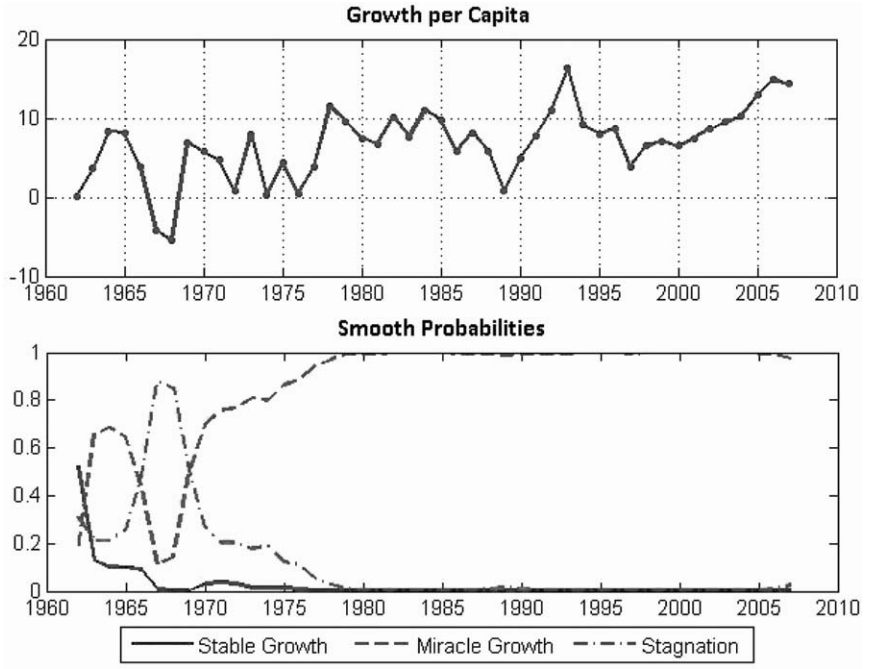


FIGURE 11. Growth and the occurrence of states for China.

1960s and 1970s, which turned the Chinese economy into something similar to a market economy, and, beginning in the 1970s, China emphasized industrialization and urban growth. Reform and liberalization are slow and continuous processes that have continued into the present day.

Figure 12 shows the regimes of Japan. In the 1960s and until the middle of the 1970s, the most common state is that of miraculous growth. Growth decreased following the oil shock, but stable growth continued. This was followed by an exhaustion of the model of high savings and high investment surpluses because Japan already had the highest per capita GDP among the industrialized countries. There was a transition toward an economy with more services and domestic consumption. Even during the "bursting of the bubble" in 1989 and the early 1990s, the model does not indicate that the country is in stagnation.

Figure 13 shows the growth pattern of the United States, which is similar to that of the countries that were already industrialized at the beginning of the sample. Stable growth is the most likely state over the entire period, with a small probability of stagnation and "catch-up" after an oil shock.

Finally, Figure 14 shows the most probable states for the Congo, which is an African country with low-quality institutions (−1.97). The stagnation state predominates, with the possibility of stable growth during the final years.

https://doi.org/10.1017/S1365100514001023 Published online by Cambridge University Press

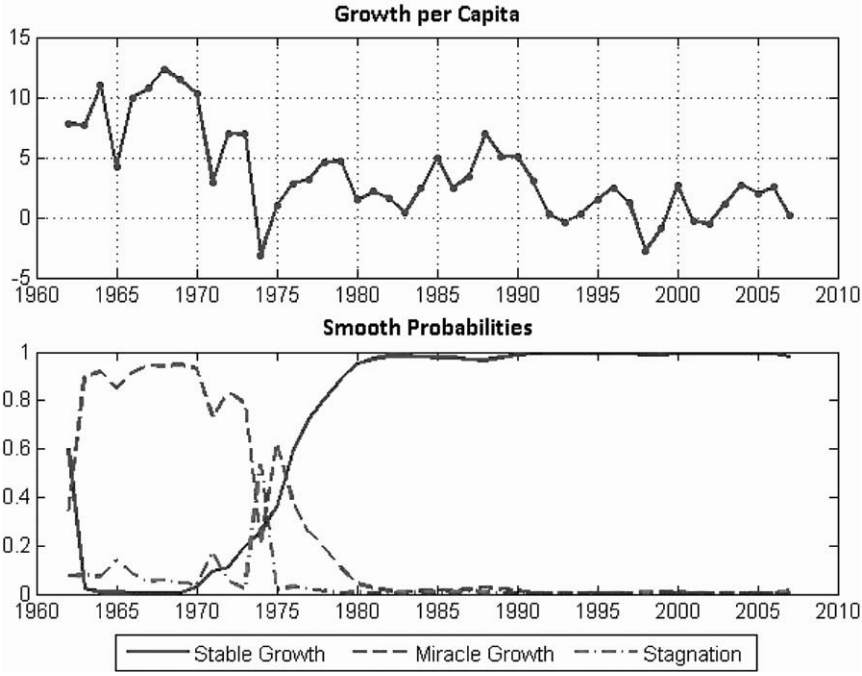


FIGURE 12. Growth and the occurrence of states for Japan.

5. CONCLUSIONS AND EXTENSIONS

In this study, we presented a model and estimation algorithm that allows different growth states with different linear submodels, in the spirit of Pritchett (2003). The transition matrix depends on the fundamental characteristics of particular countries, which change over time. With this model, it is possible to examine the variables that influence changes in the growth pattern, which is the object of study of several current studies, and even to focus on the study of growth in submodels that are appropriate for different circumstances.

In the empirical model, we develop the first aspect further, estimating a model in which the transition depends on institutional quality and on the level of investment. Three states were found that coincide with the states identified by Pritchett (2003) and Jerzmanowski (2006). The role of institutions is critical to the dynamics of growth and mainly influences recuperation from crises/stagnation and the stability of growth. An increased role for investment appears to be more restricted to countries with high-quality institutions, which increase the amount of time spent in the state of miraculous growth.

The evidence regarding the roles played by the quality of institutions and by investment suggests strong nonlinearity and conditionality for the role of the

https://doi.org/10.1017/S1365100514001023 Published online by Cambridge University Press

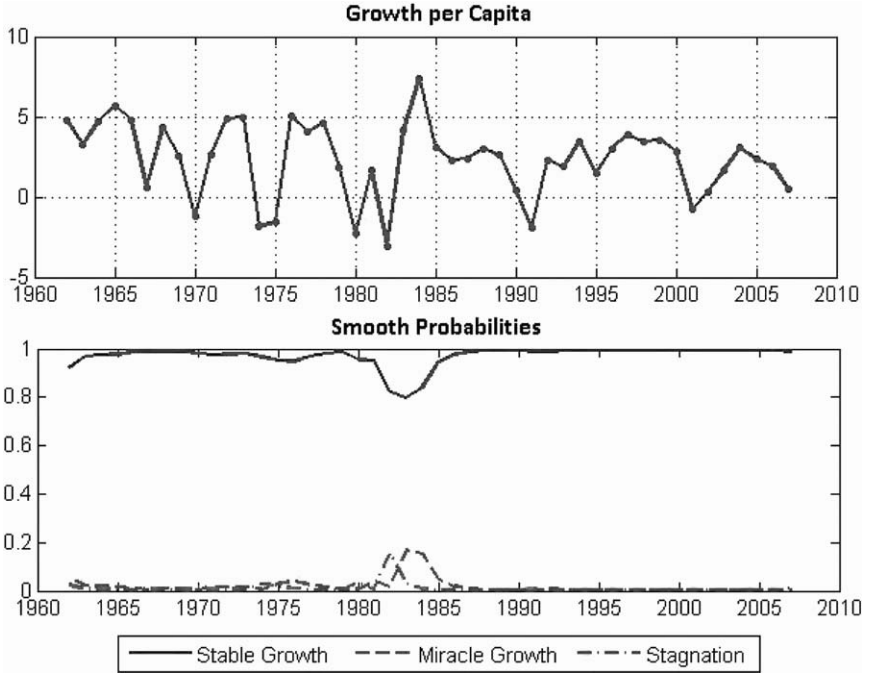


FIGURE 13. Growth and the occurrence of states for the United States.

explanatory variables, as indicated in various previous studies, such as Kourtellos (2002) Jones and Olken (2008), and Tan (2010). The empirical evidence of many countries alternating between completely different states also suggests that classical linear models are not adequate to explain growth behavior and that more complex panel models are needed to address this issue.

The estimated model also reinforces the evidence that the role of fixed effects cannot be disregarded, even in a model focusing on differences in growth. The level of the institutions (fixed for each country) significantly influences growth—which is modeled in AR(1) for each state and which is equivalent to a differential model—thus implying that the variation in growth in the estimated model is strongly conditioned by the level of growth and by the fixed effects of institutional quality.

The model developed here may be used as a more complex specification of the submodels via the incorporation of explanatory variables. We believe that this will be the most important aspect of this work. New variables, such as per capita GDP, may be incorporated into the transition matrix, as long as parsimony is maintained in the model. Eventually, by modifying the estimation algorithm, even nonlinear submodels can be considered.

https://doi.org/10.1017/S1365100514001023 Published online by Cambridge University Press

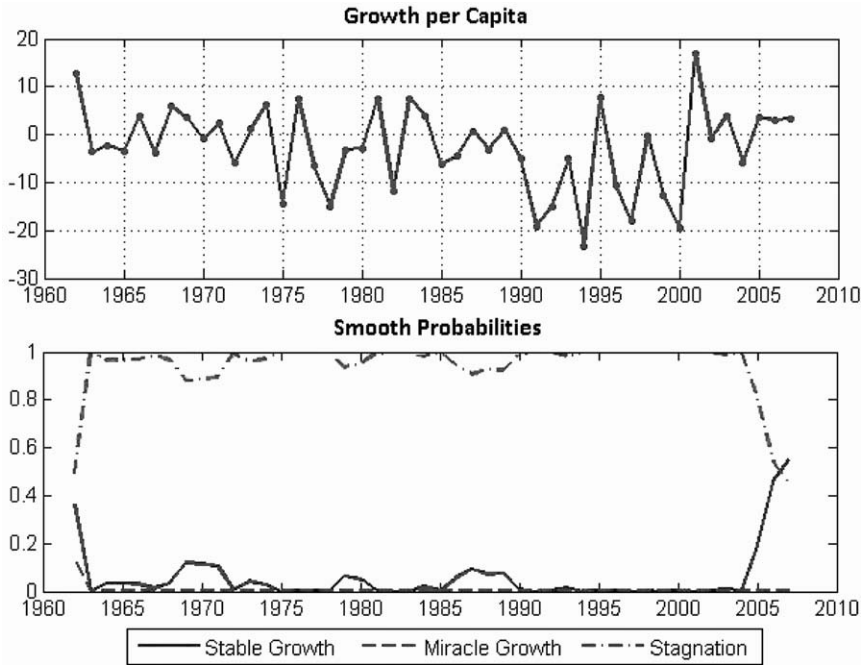


FIGURE 14. Growth and the occurrence of states for the Congo.

NOTES

1. Tan (2010) used GUIDE based on the work of Loh (2002), who extended the classification and regression tree—also known as CART—methodology from Breiman et al. (1984) and used by Durlauf and Johnson (1995).
2. Jerzmanowski (2006) used the GADP from Hall and Jones (1999).
3. Using only the data available up until time t.
4. Estimates of the probability matrix exhibit only small variations over time. Thus, although the data are annual, we understand that the estimation method is filtering short-term fluctuations, and, consequently, the use of annual data becomes appropriate.

REFERENCES

Aizenman, J. and M. Spiegel (2007) Takeoffs. Working paper 13084, National Bureau of Economic Research, Inc.

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

Ardic, O.P. (2006) The gap between the rich and the poor: Patterns of heterogeneity in the cross-country data. *Economic Modelling* 23, 538–555.

Bai, J. and P. Perron (2003) Computation and analysis of multiple structural change models. *Journal of Applied Econometrics* 18, 1–22.

Basturk, N., R. Paap, and D. van Dijk (2008) Structural Differences in Economic Growth. Discussion paper 08-085/4, Tinbergen Institute.

https://doi.org/10.1017/S1365100514001023 Published online by Cambridge University Press

Berg, A., J.D. Ostry, and J. Zettelmeyer (2012) What makes growth sustained? *Journal of Development Economics* 98, 149–166.
Bloom, D.E., D. Canning, and J. Sevilla (2003) Geography and poverty traps. *Journal of Economic Growth* 8, 355–378.
Breiman, L., J. Friedman, R. Olshen, and C. Stone (1984) *Classification and Regression Trees.* Monterey, CA: Wadsworth and Brooks.
Canova, F. (2004) Testing for convergence clubs in income per capita: A predictive density approach. *International Economic Review* 45, 49–77.
Cuberes, D. and M. Jerzmanowski (2012) Medium-term growth: the role of policies and institutions. In Olivier de La Grandville (ed.), *Frontiers of Economics and Globalization*, Vol. 11. Emerald Group Publishing Limited.
Dempster, A.P., N.M. Laird, and D.B. Rubin (1977) Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)* 39, 1–38.
Desdoigts, A. (1999) Patterns of economic development and the formation of clubs. *Journal of Economic Growth* 4, 305–330.
Diebold, F.X., J.-H. Lee, and G.C. Weinbach (1993) Regime Switching with Time-Varying Transition Probabilities. Working paper 93-12, Federal Reserve Bank of Philadelphia.
Durland, J.M. and T. McCurdy (1994) Duration-dependent transitions in a Markov model of U.S. GNP growth. *Journal of Business and Economic Statistics* 12, 279–288.
Durlauf, S.N. and P.A. Johnson (1995) Multiple regimes and cross-country growth behaviour. *Journal of Applied Econometrics* 10, 365–384.
Easterly, W., M. Kremer, L. Pritchett, and L.H. Summers (1993) Good policy or good luck? Country growth performance and temporary shocks. *Journal of Monetary Economics* 32, 459–483.
Ehrlich, I. and F.T. Lui (1999) Bureaucratic corruption and endogenous economic growth. *Journal of Political Economy* 107, S270–S293.
Filardo, A.J. (1994) Business-cycle phases and their transitional dynamics. *Journal of Business and Economic Statistics* 12, 299–308.
Hall, R.E. and C.I. Jones (1999) Why do some countries produce so much more output per worker than others? *Quarterly Journal of Economics* 114, 83–116.
Hamilton, J.D. (1990) Analysis of time series subject to changes in regime. *Journal of Econometrics* 45, 39–70.
Hamilton, J.D. (1994) *Time Series Analysis,* 1st ed. Princeton, NJ: Princeton University Press.
Hansen, B.E. (2000) Sample splitting and threshold estimation. *Econometrica* 68, 575–604.
Hausmann, R., L. Pritchett, and D. Rodrik (2004) Growth Accelerations. Working paper, John F. Kennedy School of Government, Harvard University.
Hsieh, C.-T. and P.J. Klenow (2009) Misallocation and manufacturing TFP in China and India. *Quarterly Journal of Economics* 124, 1403–1448.
Jerzmanowski, M. (2006) Empirics of hills, plateaus, mountains and plains: A Markov-switching approach to growth. *Journal of Development Economics* 81, 357–385.
Jones, B.F. and B.A. Olken (2008) The anatomy of start–stop growth. *Review of Economics and Statistics* 90, 582–587.
Jones, C.I. (2011) Intermediate goods and weak links in the theory of economic development. *American Economic Journal: Macroeconomics* 3, 1–28.
Jong-A-Pin, R. and J. de Haan (2007) Political Regime Change, Economic Reform and Growth Accelerations. CESifo Working Paper Series 1905, CESifo Group Munich.
Kerekes, M. (2009) Growth Miracles and Failures in a Markov Switching Classification Model of Growth. Discussion Paper 11, Freie Universität Berlin.
Knack, S. and P. Keefer (1995) Institutions and economic performance: Cross-country tests using alternative institutional measures. *Economics and Politics* 7, 207–227.
Kourtellos, A. (2002) A Projection Pursuit Approach to Cross Country Growth Data. Working papers in economics 0213, University of Cyprus Department of Economics.

https://doi.org/10.1017/S1365100514001023 Published online by Cambridge University Press

Loh, W.-Y. (2002) Regression Trees with Unbiased Variable Selection and Interaction Detection. Technical report, Statistica Sinica.

Masanjala, W.H. and C. Papageorgiou (2004) The Solow model with CES technology: Nonlinearities and parameter heterogeneity. *Journal of Applied Econometrics* 19, 171–201.

Paap, R., P.H. Franses, and D. van Dijk (2005) Does Africa grow slower than Asia, Latin America and the Middle East? Evidence from a new data-based classification method. *Journal of Development Economics* 77, 553–570.

Pritchett, L. (2000) Understanding patterns of economic growth: Searching for hills among plateaus, mountains, and plains. *World Bank Economic Review* 14, 221–250.

Pritchett, L. (2003) A toy collection, a socialist star and a democratic dud: Growth theory, Vietnam, and the Philippines. In Dani Rodrik (ed.), *In Search of Prosperity: Analytical Narratives on Economic Growth*. Princeton, NJ: Princeton University Press.

Ruud, P.A. (1991) Extensions of estimation methods using the EM algorithm. *Journal of Econometrics* 49, 305–341.

Tan, C.M. (2010) No one true path: Uncovering the interplay between geography, institutions, and fractionalization in economic development. *Journal of Applied Econometrics* 25, 1100–1127.

Wacziarg, R. and K.H. Welch (2003) Trade Liberalization and Growth: New Evidence. Research paper, Graduate School of Business, Stanford University.

Watson, M. (1983) Alternative algorithms for the estimation of dynamic factor, mimic and varying coefficient regression models. *Journal of Econometrics* 23, 385–400.

https://doi.org/10.1017/S1365100514001023 Published online by Cambridge University Press