Escola de ECONOMIA de São Paulo

A TIME-VARYING MARKOV-SWITCHING MODEL FOR ECONOMIC GROWTH

BRUNO MORIER
VLADIMIR KUHL TELES

FGV

EESP

A Time-Varying Markov-Switching Model for Economic Growth

Bruno Morier[*]

Vladimir Kühl Teles[†]

November 11, 2011

Abstract

This paper investigates economic growth's pattern of variation across and within countries using a Time-Varying Transition Matrix Markov-Switching Approach. The model developed follows the approach of Pritchett (2003) and explains the dynamics of growth based on a collection of different states, each of which has a sub-model and a growth pattern, by which countries oscillate over time. 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 the Diebold's EM Algorithm and estimate an example model in a panel with a transition matrix conditioned on the quality of the institutions and the level of investment. We found three states of growth: stable growth, miraculous growth, and stagnation. The results show that the quality of the institutions is an important determinant of long-term growth, whereas the level of investment has varying roles in that it contributes positively in countries with high-quality institutions but is of little relevance in countries with medium- or poor-quality institutions.

Key-Words: Growth Regimes, Markov-Switching.
JEL Class: R14, H0.

---

[*]EESP/FGV and Banco Itaú
[†]EESP/FGV

1 Introduction

Economic growth varies among different countries and over time. Several articles attempt to explain the differences in cross-country growth rates, showing that countries tend to converge to different growth regimes (Desdoigts, 1999; Durlauf and Johnson, 1995; Masanjala and Papageorgiou, 2004; Tan, 2010). Other articles show that a country can move between regimes of growth, alternating growth "miracles" and "failures" (Easterly et al., 1993; Hausmann et al., 2004; Jerzmanowski, 2006; Jones and Olken, 2008; Pritchett, 2000, 2003). This paper proposes a methodology to study the differences in growth rates of countries as well as the growth process changes over time employing a Markov regime-switching approach. However, instead of considering that some countries have been fortunate to have a "good transition probability matrix" and others are born with a "bad transition matrix", as is usually done, we develop a generalization of the Diebold's EM Algorithm and estimate a Time-Varying Transition Matrix Markov-Switching model. Thus we estimate the parameters of the transition matrix varying over time, conditioned on the quality of the institutions and the level of investment.

The analyses performed in most studies do not incorporate all available information related to growth and they do not consider the intertemporal factor. More recently, however, there have been several studies that focus on precisely this aspect by dating regime changes and finding explanatory variables. In Jones and Olken (2008), the methodology of Bai and Perron (2003) was used to date breaks in the growth process, and a descriptive panel with means and standard deviations was included for changes in variables correlated with growth, which was incorporated into a Solow residual model that described the anatomy of the variations in growth. In Hausmann et al. (2004), the break for a better growth pattern (acceleration) is analyzed. One ad-hoc criterion for acceleration was selected and calculations were performed using means and standard deviations for the periods prior to and following the acceleration event. In Wacziarg and Welch (2003), a similar study was performed that investigated the average impact of the removal of restrictions on external trade on economic growth. Other studies using the same approach include Aizenman and Spiegel (2007) and Jong-A-Pin and de Haan (2007). 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 large degree of homogeneity in the production function among different countries, where this is always identical and in the Cobb-Douglas format. The differences in growth patterns among countries are obvious, and the production functions may be quite heterogeneous, which suggests the need for a more flexible approach. In Durlauf and Johnson (1995), the case of a linear and homogenous Cobb-Douglas function was rejected

in favor of a model with multiple growth regimes. In contrast to the approach in our study, the regimes were fixed for each country, according to the initial conditions, from the beginning to the end of the sample. This approach generated a series of literature on convergence clubs and growth regimes.

Durlauf and Johnson (1995) employed the Regression Tree technique to identify the regimes, and other studies suggest 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 for tests. Masanjala and Papageorgiou (2004) have used it to perform these tests, finding non-linearity of the production function and a constant elasticity substitution (CES) format, which would be a possible explanation for the heterogeneity of the coefficients and the existence of multiple regimes. In Durlauf and Johnson (1995), four regimes were found. In Desdoigts (1999) and in Kourtellos (2002), the projection pursuit technique is used to approximate the growth functions by projections on lower dimension functions. In Alfo et al. (2008); Bloom et al. (2003); Paap et al. (2005) and Basturk et al. (2008), the regime to which a country belongs is estimated along with the other parameters using mixture analysis with classical statistics, and in Ardic (2006), the Bayesian statistic is used. In Canova (2004), the predictive density technique was used. In Tan (2010), the regression tree technique is revisited with the use of Generalized Unbiased Interaction Detection and Estimation (GUIDE)[1].

The fixed regime approach, however, does not explain the inconsistencies in growth or breaks in regimes that are characterized by the experiences of developing countries and that strongly contrast with the experiences of the most advanced industrialized countries upon which the classical theories of economic growth are based. In Pritchett (2000), this inconsistency is examined among the different countries, and six different growth patterns are identified to analyze the dynamic of the growth level along with breaks and variations. In Pritchett (2003), the use of an approach with six different states of growth through which countries pass was suggested. Each state is associated with a growth level that is determined by a different sub-model, and the growth process is then explained by toy-collection models. Thus, the experience of inconsistency and breaks of the developing countries is explained by changes in state, which are less common in the case of developed countries. With this approach, multiple equilibria are permitted, such as poverty traps, heterogeneous production functions, and different roles for the conditioning variables, which are non-linear and have breaks. These factors are supported by strong empirical evidence. The changes in state are modeled probabilistically with the probabilities conditioned by the prevailing foundations of each country. In an experiment with a constant ad-hoc transition matrix, the author managed to recreate the dispersion of the current per capita GDP levels.

---

[1]He used GUIDE(*Generalized Unbiased Interaction Detection And Estimation*) based on the work of Loh (2002) who extended the Classification and Regression Tree (CART) methodology (*Classification and Regression Tree*) from Breiman et al. (1984) andused by Durlauf and Johnson (1995)

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

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

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 asymmetrical and non- linear (i.e., depending on the current state or influencing the transition among states)

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 the model that is 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 an example model in which the transition matrix was conditioned by the quality of the institutions and the level of investment. The primary results indicated that the quality of the institution is an important determinant of long-term growth, whereas the level of investment plays a differentiated role: it positively contributes towards growth in countries with high-quality institutions but is of little relevance in countries with medium- or poor-quality institutions.

---

[2]He used the GADP from Hall and Jones (1999)

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 and opens 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 time $t-1$ and t. g_{ct} is modeled following a Time-Varying Markov-switching process with S AR(p) states. The precise meaning is that for each time t that 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 $z^{ct} = (z_i^{ct})_{i=1...N_z}$ 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 even though the value of these variables varies with c and t. For each state, the AR(p) process is given by the expression:

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

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 is conditional on exogenous explanatory variables. These N_x conditional variables are collected in the vector $x^{ct} = (x_i^{ct})_{i=1...N_x}$ 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 $p_{ij}(x^{ct})$:

$$
p_{ij}(x^{ct}) = \begin{cases} \dfrac{\exp(\beta_{ij}x^{ct})}{1 + \displaystyle\sum_{k=0}^{S-2}\exp(\beta_{kj}x^{ct})} & , \text{if } i = 0, \cdots, S-2 \\[4ex] 1 - \displaystyle\sum_{k=0}^{S-2} p_{kj}(x^{ct}) & , \text{if } i = S-1 \end{cases}
\tag{3}
$$

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 a vector ($N_x \times 1$).

Note that this is a flexible model, with the explanatory variables conditioning the process of each state and the transition matrix. It is more realistic than the usual Markov-switching specification with a fixed matrix (as in Hamilton (1994)), as 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 good institutions and good macroeconomic policies is less than that in a country with poor institutions and policies. The fundamental characteristic may change over time, as is the case with investment and the political regime.

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

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

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

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

Maintaining the transition matrix P_{ct} 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. Using $\pi = (\pi_0, \cdots, \pi_{S-1})'$ as the unconditional probability, equation (4) implies $\pi = P_{ct}\pi$. We also have $\sum_i \pi_i = 1$. As in Hamilton (1994), the solution to this system is given by

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

6

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

$$
A \atop (S+1 \text{ X } S)' = \left[\begin{array}{c} I_S - P_{ct} \\ 1 \cdots 1 \end{array} \right] \tag{7}
$$

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 previous 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). We can cite other Markov-switching algorithms with EM: Hamilton (1990), Durland and McCurdy (1994), and Filardo (1994).

Following the model described in the previous section, the data available for the estimation should be g_{ct}, z^{ct}, and x^{ct}. We assume that these data are available 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 $\underline{x_{cT}} = (x_{c1}, \ldots, x^{ct})$. Thus, the input for the estimation is $\underline{g_{cT}}, \underline{x_{cT}}, \underline{z_{cT}}$.

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 $\theta = \left(\alpha, \beta, \gamma, \sigma, \rho \right)$. These 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 estimations for the parameters of the model. Even though this is not the case, the calculation of this function proves to be 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). As 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 a lighter notation, we will remove c, using g_t instead of g_{ct}. We then have

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

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

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

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

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

where $\rho_j = P(s_1 = j)$ is the j-th 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

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

We observe that the operator, $E[.]$, essentially substitutes $I(.)$ for $P(.)$ in the equation. Using the independence, the function of log maximum likelihood for the panel of countries is obtained by the sum

8

of the functions of log maximum likelihood for each country. Let Ψ be this function and ψ_c be the log function of maximum likelihood for each country. We then have

$$
\begin{aligned}
\Psi = \sum_c \Psi_c = & \sum_c \sum_{j=0}^{S-1} \rho_j \big(\log f(g_{c1}|s_{c1} = j, z_{c1}, \theta) + \log \rho_{cj} \big) \\
+ & \sum_c \sum_{t=2}^{T} \sum_{j_1=0}^{S-1} \sum_{j_2=0}^{S-1} P(s_{ct} = j_1, s_{ct-1} = j_2 | \underline{g}_{c_T}, \underline{x}_{c_T}, \underline{z}_{c_T}, \theta) \\
\times & \ \left\{ \log f(g_{ct}|s_{ct} = j_1, \underline{g}_{c_{t-1}}, z^{ct}, \theta) + \log P(s_{ct} = j_1|s_{ct-1} = j_2, x^{ct}, \theta) \right\}
\end{aligned}
\tag{13}
$$

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

3.1 EM Algorithm - General Procedure

Given the model from Section 2, we have created formulae for all of 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 occurrence of states $s_{ct} = j_1, s_{ct-1} = j_2$ that is conditional on all of the data previous and subsequent to time t. We can obtain this term algorithmically. Because this term is conditional on all of the data, this term tends to have variations that are lower than those of the other terms of the equation for values of θ that are close to the maximum function of expected maximum likelihood.

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

(1) Select a $\theta^{(0)}$

(2) With the last $\theta^{(l)}$ available, calculate

$\quad P(s_t = j_1, s_{t-1} = j_2 | \underline{g}_T, \underline{x}_T, \underline{z}_T, \theta^{(l)})$ para $(j_1, j_2) \in (0, \ldots, S-1)^2$ e $t \in (2, \ldots, T)$

(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 $\theta^{(l)}$ available, i.e., by substituting $P(s_{ct} = j_1, s_{ct-1} = j_2 | \underline{g}_{c_T}, \underline{x}_{c_T}, \underline{z}_{c_T}, \theta)$ by $P(s_{ct} = j_1, s_{ct-1} = j_2 | \underline{g}_{c_T}, \underline{x}_{c_T}, \underline{z}_{c_T}, \theta^{(l)})$ 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 numeric approximation of the maximum point 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 as described in Diebold et al. (1993). Again, we will treat each country separately, removing c from the notation.

1. Pre-calculus

   (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

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

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

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

   (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 $P(s_t, s_{t-1}|\underline{g}_T, \underline{x}_T, \underline{z}_T, \theta^{(l)})$ to $(s_t, s_{t-1}) \in (0, \ldots, S-1)^2$ and $t \in (2, \ldots, T)$

---

[3] using only the data available up until time 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

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

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

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

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

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

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

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

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

$$
\xi_j^{ct}=
\begin{cases}
\displaystyle\sum_{j_2=0}^{S-1}\xi_{j_1,j_2}^{ct}\,, \text{ if } t>1 \\[4mm]
\rho_j^c\,, \text{ if } t=1
\end{cases}
\tag{22}
$$

Applying $\log$ in equation (10), we obtain

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

Using equation (22) and the above definitions (21), we have a simpler representation for Ψ^l derived from equation (13), and the substitution of the dependent smoothed probabilities on θ by $\theta^{(l)}$.

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

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

1. Calculation of α and of γ

Deriving equation (23), we have

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

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

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

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

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

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 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 resolving this system.

2. Calculation of σ

Deriving equation (23), we obtain

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

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

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

The volatilities of each state may then be calculated with

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

3. <u>Calculation of β</u>

   Deriving 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

   $$1 - \sum_{k=0}^{S-2} p_{k,j_2}^{ct} = \left(1 + \sum_{k=0}^{S-2} \exp(\beta_{k,j_2} x^{ct}) \right)^{-1}$$

   $$\tag{35}$$

   $$\frac{\partial}{\partial \beta_{i_1,i_2}} \log \left(1 + \sum_{k=0}^{S-2} \exp(\beta_{k,i_2} x^{ct}) \right) = x^{ct} p_{i_1,i_2}^{ct}$$

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

   $$\sum_{j_1=0}^{S-1} \xi_{j_1,j_2}^{ct} \log p_{j_1,j_2}^{ct} = \sum_{j_1=0}^{S-2} \xi_{j_1,j_2}^{ct} \log \left(\frac{\exp(\beta_{j_1,j_2} x^{ct})}{1 + \sum\limits_{k=0}^{S-1} \exp(\beta_{k,j_2} x^{ct})} \right) + \xi_{S-1,j_2}^{ct} \log \left(1 - \sum_{k=0}^{S-2} p_{k,j_2}^{ct} \right)$$

   $$= \sum_{j_1=0}^{S-2} \xi_{j_1,j_2}^{ct} \beta_{j_1,j_2} x^{ct} - \xi_{j_2}^{ct} \log \left(1 + \sum_{k=0}^{S-2} \exp(\beta_{k,j_2} x^{ct}) \right) \tag{36}$$

   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 especially good as we approach the convergence criterion, i.e., the real β.

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

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

$$\phi_{i1,i2}^{ct} = \frac{\partial p_{i_1,i_2}^{ct}}{\partial \beta_{i_1,i_2}}(\beta^{(l)}) = \frac{\partial}{\partial \beta_{i_1,i_2}} \left[\frac{\exp(\beta_{i_1,i_2} x^{ct})}{1 + \sum_{k=0}^{S-1} \exp(\beta_{k,i_2} x^{ct})} \right] = (x^{ct})' \left[p_{i_1,i_2}^{ct}(\beta^{(l)}) - p_{i_1,i_2}^{ct}(\beta^{(l)})^2 \right] \tag{39}$$

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

$$\beta_{i_1,i_2} = \beta_{i_1,i_2}^{(l)} + \left[\sum_{c,t} x^{ct} \xi_{i_2}^{ct-1} \phi_{i_1,i_2}^{ct} \right]^{-1} \left[\sum_{c,t} x^{ct} \left(\xi_{i_1,i_2}^{ct} - \xi_{i_2}^{ct-1} p_{i_1,i_2}^{ct}(\beta^{(l)}) \right) \right] \tag{40}$$

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 by the quality of the institutions and by the level of investment (%GDP). Each state is modeled following an AR(1) without explanatory exogenous variables.

The resulting model possesses three states, which is one less than is found in Jerzmanowski (2006). We are unaware of the objective statistical criterion in the literature for the selection of the number of states in Markov-switching models similar to the one used in this work. However, the selection is made 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; i.e., its resulting transition matrix presents many cases of 100% or 0% transition probability, indicating few occurrences of transition conditional on the effective values for the pair of explanatory variables.

Despite having one less state, the results corroborate those of the model of Jerzmanowski (2006) 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 work is actually 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 for the result of Jerzmanowski (2006), as the results are analogous to the introduction of investment with a different data set: the author's data were from 89 countries during a period from 1962 to 1994 with a Penn World Tables 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 appears to be equivalent to the union of "steady growth at levels of income near world leaders" and the "non-converging steady growth" state. 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 between 1962 and 2007 were considered[4]

To measure the quality of the institutions, 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 the bureaucracy, among others. The index is assumed to be constant over the time period, varying only by country, as in Jerzmanowski (2006). This variable will be denoted as GADP_c.

To measure the level of investment, we use the 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 $(t-5, \ldots, t-1)$. For some countries in the sample, an extrapolation for the initial years was performed due to the lack of data. However, the GFCF series exhibits slow changes during this time, and even considering that the series used is the average of 5 years, we should not be biased by the procedure that was adopted. From this point on, this variable (average GFCF of the 5 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]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, so the use of annual data becomes appropriate.

4.2 Results

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

Table 1: States: AR(1) Process

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

Table 1 presents the parameters of the AR(1) process described in 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 (5).

State 1, which we call stable growth, is the state that best represents the experience of the developed countries with a growth rate close to 2%, a high auto-correlation among the growth rates, and low volatility. Countries such as Colombia and Pakistan with stable and non-convergent 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 historical data that are available, the countries that maintained this state for the longest period of time are the emerging Asian countries, with illustrative cases such as Japan and China.

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

Tables 2,3,4 present the most important examples of countries and periods of incidence for each of the states. The incidence of each state in these tables follows the convention with a probability greater than 75% for the occurrence of the state. The variables of GADP and GFCF are included, with values that are normalized to the mean and standard deviation for the entire sample, including 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 criterion for the formation of table 2 was countries with 30 years or more in state 1, summing all episodes. We observe in table 2 that the mean of the normalized GADP is high, i.e., 0.72, indicating

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 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 &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				

a relationship between this variable and the occurrence of stable growth. In addition, most countries that do not have a high GADP have episodes where 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 non-convergent growth" by Pritchett (2003). This also appears to be the case for the 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 countries with 10 years or more in state 2, summing all episodes. We observe in table 3 that the mean GFCF is high, i.e., 0.96, indicating a relationship between this variable and the occurrence of miraculous growth. In addition, most of the countries that do not possess a high GFCF maintain miraculous growth for a few years. The Asian countries tend to maintain this type of growth. Greece decreased its rate of investment after the occurrence of this state. Japan reached convergence of its per capita GDP with that of the leading countries, concluding the "catch-up" phase and beginning 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.

Table 4 shows countries with 20 years or more in state 3, summing all episodes. We observe in table 4 that the mean GFCF and mean GADP are low, at -0.72 and -0.47, marking a relationship between these variables and the occurrence of stagnation.

### 4.2.1 The Role of the Quality of the Institutions and of Investment

In this subsection, we explain the relationships among the quality of the institutions, investment, and the transition matrix. As discussed in section 1, the model that has been adopted allows for a more

Figure 1: The Impact of Investment for Countries with a High Quality of Institutions
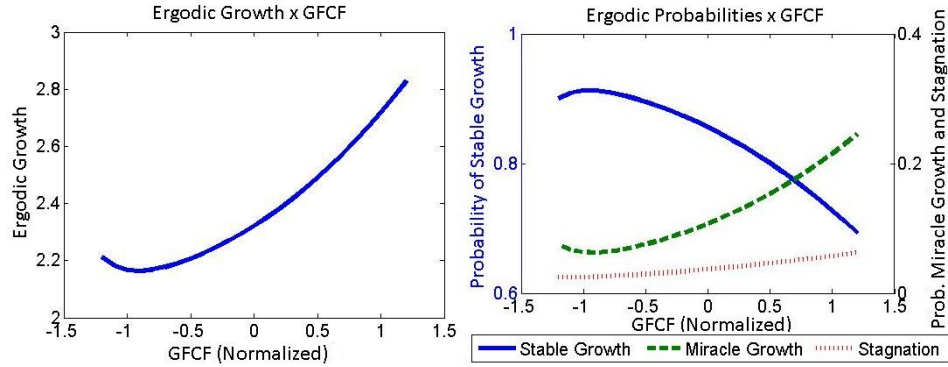
Note: Normalized GADP set at 1.2.

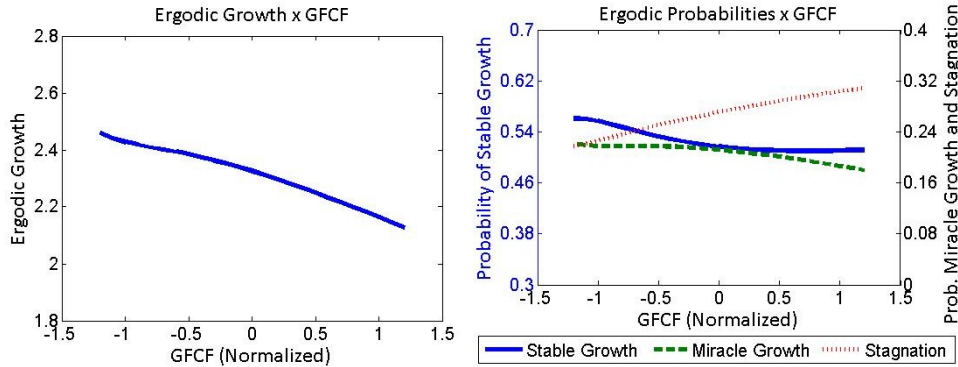Figure 2: The Impact of Investment for Countries with a High Quality of Institutions
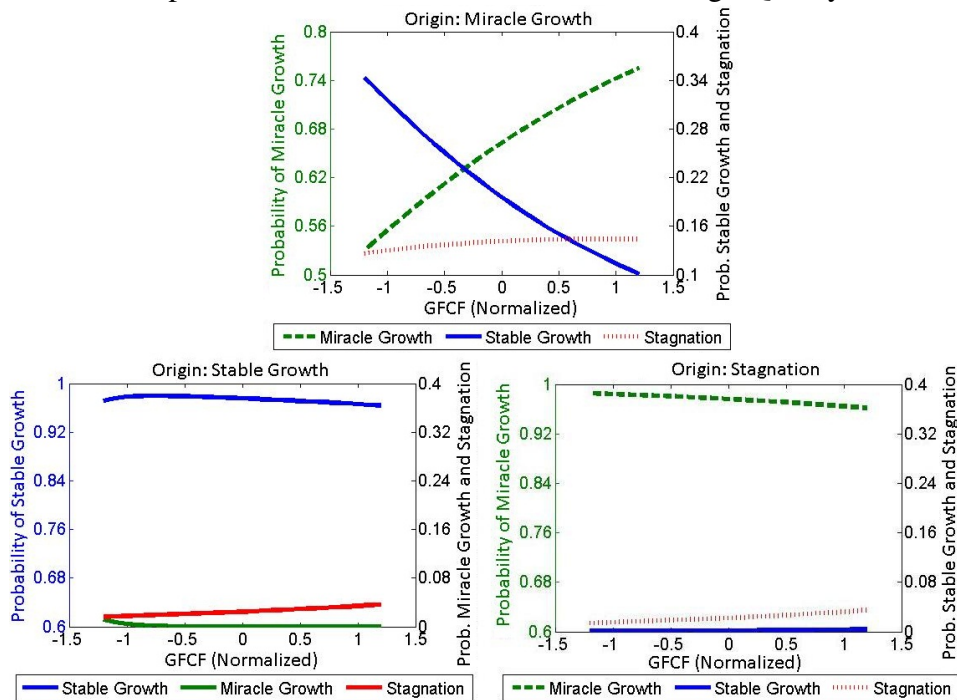
Note: Normalized GADP set at 1.2.

complete description of the impact of each variable than does a purely linear model. As two variables are under consideration, we first normalized the level of the institutions in three levels $(-1.2, 0, 1.2)$, and we analyzed the impact of investment for each one of these levels. The levels were selected for observation of the dynamics not only for countries with average institutions but also for those with good or poor institutions. The selected levels correspond approximately to the 10% and 90% percentiles. Next, we established the same three levels for investment for a similar analysis.

In figure 1, we observe the impact of investment for countries with high-quality institutions. The graph on the left shows the ergodic growth (long-term) and the one on the right shows the ergodic probabilities (unconditional) of occurrence for each state. These are calculated from equations (5), (6) and (8). We observe that there is a positive impact on investment: the highest levels of investment result in a greater

Figure 3: The Impact of Investment for Countries with Average-Quality Institutions

Note: Normalized GADP set at 0.

Figure 4: The Impact of Investment for Countries with Low-Quality Institutions
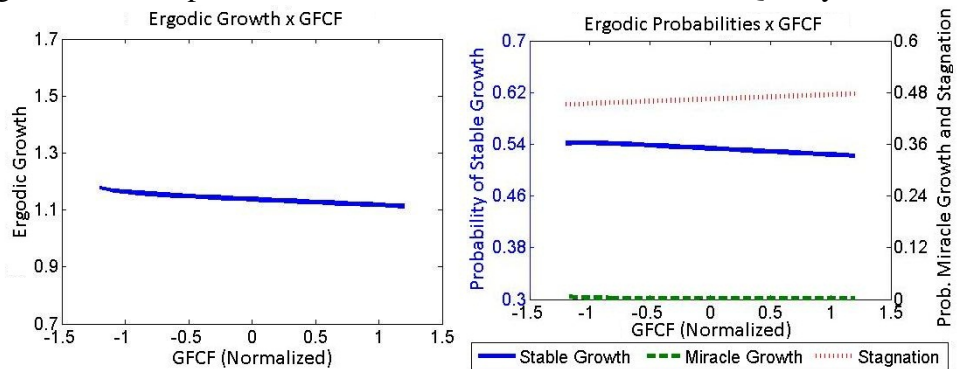
Note: Normalized GADP set at -1.2.

amount of time allocated to the miraculous growth state, thus generating greater long-term growth.

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

Figure 3 shows the case of medium-quality institutions. In this case, an increase in investment does not significantly impact growth. The effect even appears to be slightly negative due to the amount of time in stagnation. However, the effect appears to be somewhat irrelevant, generating a variation of 0.3% from end to end in the level of investment. We also observe that, for this level of institution quality, miraculous growth is frequent and is counterbalanced in terms of the average growth by the stagnations, which are also frequent. The growth is then 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 the income level, as was predominant in the previous case.

Figure 5: The Impact of Institutions for High-Investment Countries
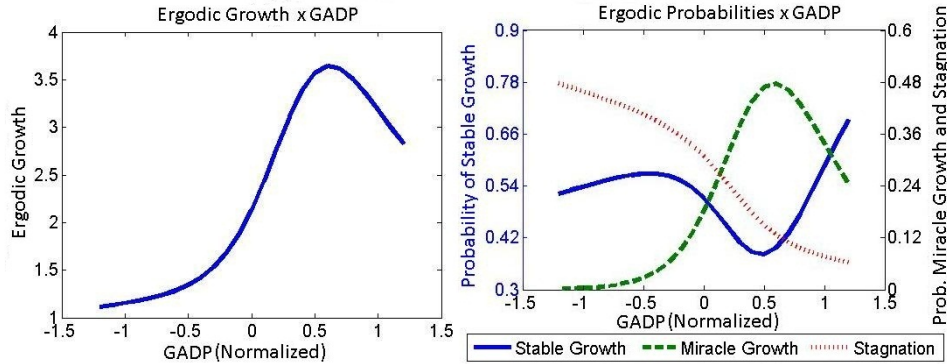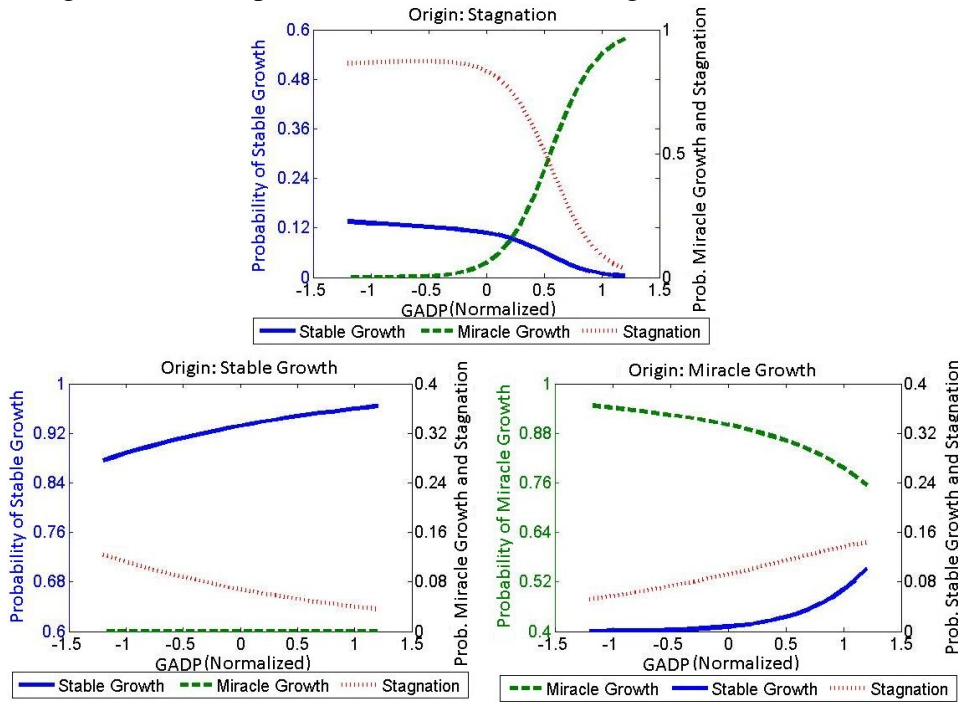
Note: Normalized GFCF set at 1.2.

Lastly, in figure 4, we have the case of below-average institution quality. In this case, the growth dynamics are not favorable, and the level of investment does not have a positive effect. The effect is irrelevant, is less than in the previous case, and even tends to be negative. We continue with the analysis of a fixed level of investment. In figure 5, we have a dynamic for a high level of investment. We observe that the ergodic growth is increasing in most of the graph and that the institutions exert a strong impact over growth. In general, better institutions increase the occurrence of miraculous growth and decrease stagnation.

The exception is found for the final plot above 0.7 in the level of institution quality, where there is a decline in miraculous growth. This effect, which also occurs in the model of Jerzmanowski (2006), is given by the correlation between the high quality of the institutions and the per capita GDP level. In the sample, there is a concentration of countries with a high per capita GDP level among the countries with institutions of very high quality, and this high per capita GDP level decreases the incidence of miraculous growth. According to Pritchett (2003), this is due to the "catch-up" in growth convergence.

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

In figure 6, we have a transition matrix with a high level of investment for the function of the quality of the institutions. We can find the reasons for a lesser amount of time in stagnation in the first two tables. When in stagnation, the economy with better institutions has an increased chance of recovery and moving into miraculous growth, whereas the countries with poor-quality institutions have a high probability of remaining stagnant, i.e., as in a poverty trap. When in a state of steady growth, the economy with better

22

Figure 6: The Impact of the Institutions for High-Investment Countries

Note: Normalized GFCF set at 1.2.

Figure 7: The Impact of Institutions for Medium-Investment Countries
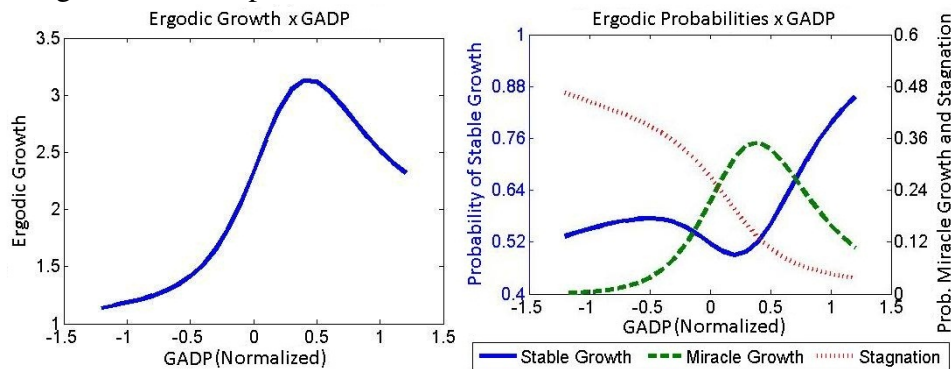
Note: Normalized GFCF set at 0.

institutions has greater stability, whereas the economy with worse institutions has a greater chance of falling into stagnation, even though there is 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, segueing into stable growth. We observe that the increase in the probability of stagnation is minimal and immaterial as the incidence of stagnation in the highest levels of institution quality is not perennial, and it is more likely that miraculous growth will occur, where this will occur even more quickly as the level of the institutions increases.

In figures 7 and 8, we have the case 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.

Figure 8: The Impact of Investment for Low-Investment Countries
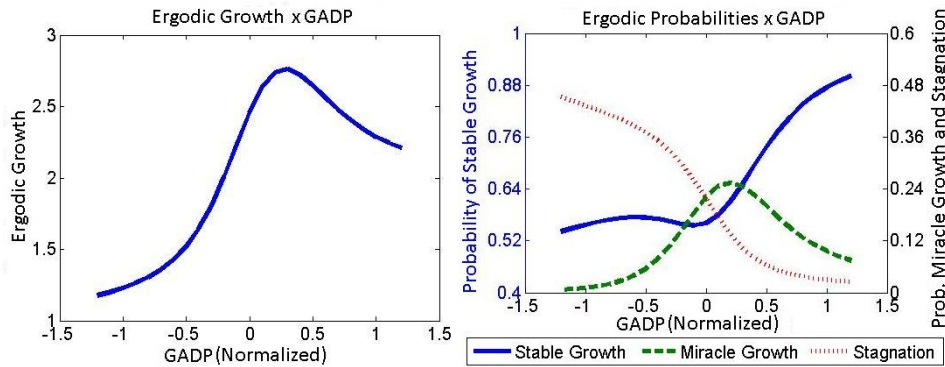
Note: Normalized GFCF set at -1.2.

Figure 9: Growth and the Occurrence of the States for Brazil
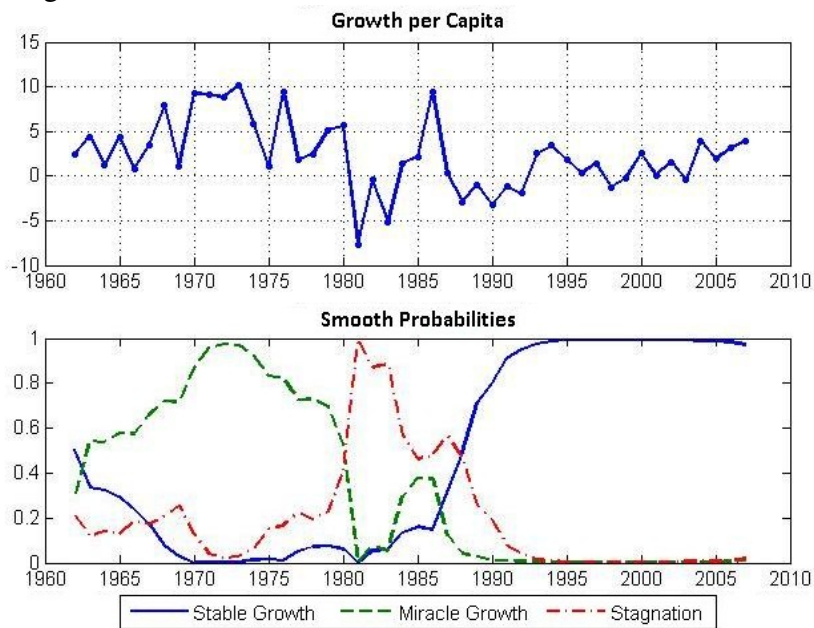
4.2.2 Cases of Selected Countries

In this section, we present the states that are attributed by the model to six selected countries: Brazil, Argentina, China, Japan, the United States, and the Congo. The objective is to use practical cases to indicate how the change of regimes works and to show that certain factors may impact these changes. The model seeks to explain the probabilities of change conditional to the structural variables, but most of the changes have cyclical and specific aspects that lead to the transition.

In figure 9, we examine the case of Brazil. In the 1960s and the 1970s, we have a dominant probability of miraculous growth. Despite the rising debt and the troubled political and social scenes that characterized this period, there was strong growth spurred by investment, particularly 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 crisis, the country had to implement adjustments for the previous

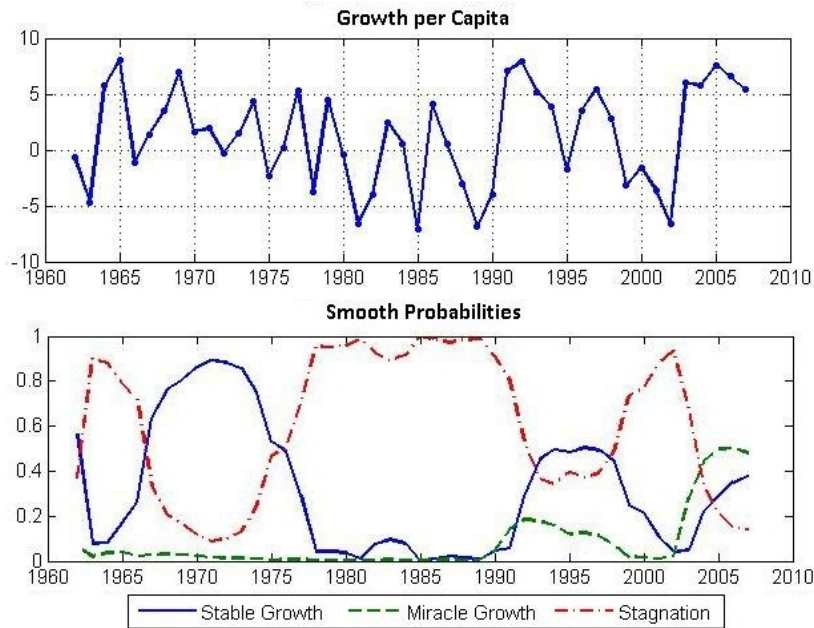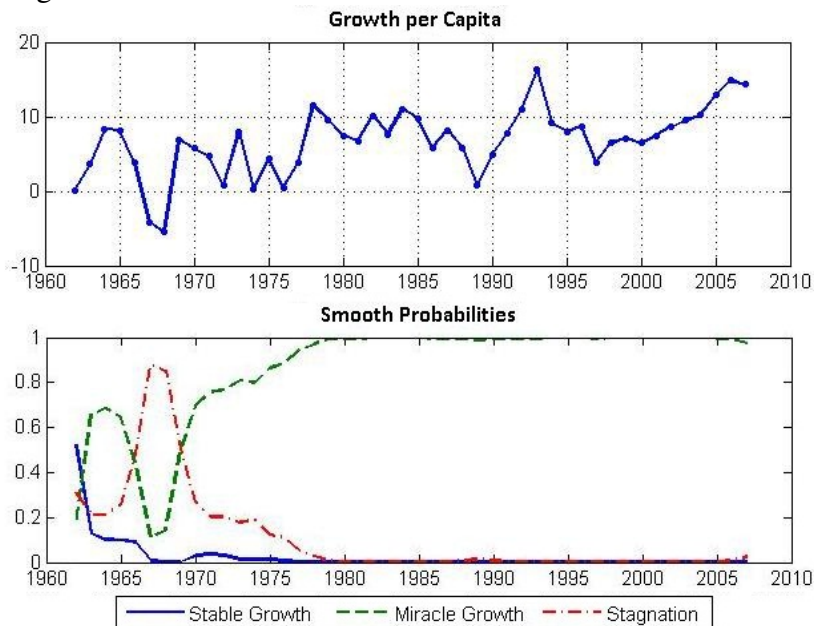Figure 10: Growth and the Occurrence of the States for Argentina

Figure 11: Growth and the Occurrence of the States for China

excesses in the external and fiscal accounts and, furthermore, had to control for inflation; all of this occurred at a time of transition in the political regime with the end of the dictatorship. After the constitution of 1988 and the control of inflation, there was stabilization of growth at a non-convergent stable state, to apply the terminology of Pritchett (2003).

Figure 10 shows the Argentinean case. According to Pritchett (2003), Argentina, for the most part, faces a situation of stagnation, but this 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

Figure 12: Growth and the Occurrence of the States for Japan
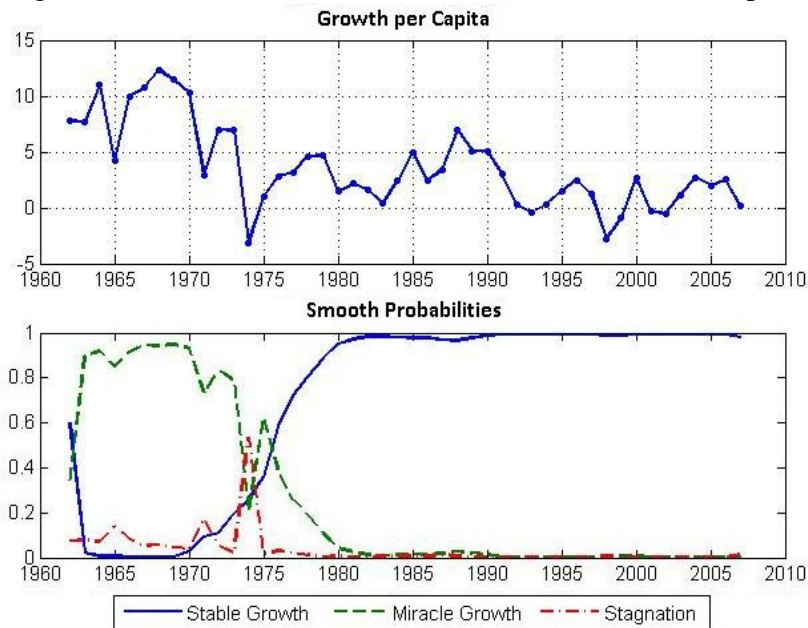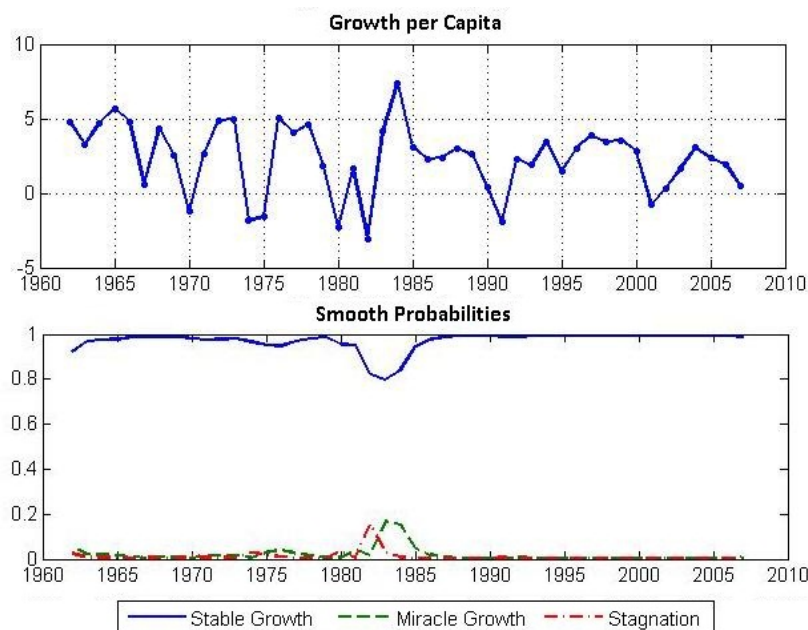

Figure 13: Growth and the Occurrence of the States for the United States

some degree of recuperation during the reforms of the Menem era, there was a new crisis period with an external deficit, capital flight, and non-payment of debt.

In figure 11, we present the impressive case of China. With the greatest level of investment, as a % of the GDP, China has maintained miraculous growth since the 1970s. The most rapid growth process began with the reforms during the 1960s and 1970s, which turned the Chinese economy into something similar to a market economy and emphasized industrialization and urban growth beginning in the 1970s. The reforms and the liberalization are slow and continuous processes that have continued until the present

Figure 14: Growth and the Occurrence of the States for the Congo
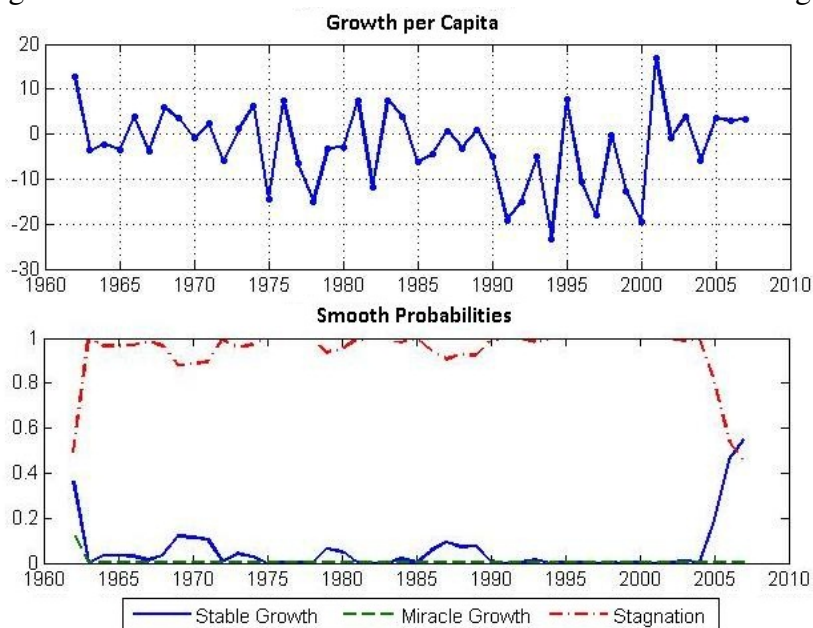
day.

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

Figure 13 shows the growth pattern of the United States, which is similar to that of the countries that were already industrialized in 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" in the event of an oil shock.

Lastly, figure 14 shows the most probable states for the Congo, which is an African country with below average institutions (-1.97). The stagnation state is predominant, with a possibility of stable growth during the final years.

5 Conclusions and Extensions

In this study, we presented a model and estimation algorithm that allow for different growth states with different linear sub-models, in the spirit of Pritchett (2003). The transition matrix depends on the fundamental characteristics of each country, which change over time. With this model, it is possible to examine the variables that influence the changes in the growth pattern, which is the object of study of various cur-

rent works, and even to specialize in the study of growth in sub-models that are appropriate for different circumstances.

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

The evidence of the roles played by the quality of the institutions and by investment suggest strong non-linearity and conditionality in the role of the explanatory variables, as indicated in various previous studies, such as Kourtellos (2002), Tan (2010), and Jones and Olken (2008). The alternation between states suggests that the series has many ruptures when analyzed by a linear framework, indicating the need for more complex models for panel studies.

The estimated model also reinforces the evidence that the role of fixed effects cannot be disregarded, even in a model for 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, implying that the variation of growth in the estimated model is strongly conditioned by the level of growth and by the fixed effects of the quality of the institutions.

The model that has been presented may be used as a more complex specification of the sub-models, incorporating explanatory variables. We believe that this will be the most important sequence of this work. New variables, such as per capita GDP, may be incorporated into the transition matrix, taking care to maintain parsimony in the model. Eventually, with a change of the estimation algorithm, even non-linear sub-models can be considered.

References

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

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

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

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

Basturk, N., Paap, R. and van Dijk, D. (2008), Structural differences in economic growth, Tinbergen Institute Discussion Papers 08-085/4, Tinbergen Institute.

Bloom, D. E., Canning, D. and Sevilla, J. (2003), "Geography and poverty traps", *Journal of Economic Growth* , Vol. 8, pp. 355–78.

Breiman, L., Friedman, J., Olshen, R. and Stone, C. (1984), *Classification and Regression Trees*, Wadsworth and Brooks, Monterey, CA.

Canova, F. (2004), "Testing for convergence clubs in income per capita: A predictive density approach", *International Economic Review* , Vol. 45, pp. 49–77.

Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977), "Maximum likelihood from incomplete data via the em algorithm", *Journal of the Royal Statistical Society. Series B (Methodological)* , Vol. 39, Blackwell Publishing for the Royal Statistical Society, pp. 1–38.

Desdoigts, A. (1999), "Patterns of economic development and the formation of clubs", *Journal of Economic Growth* , Vol. 4, pp. 305–30.

Diebold, F. X., Lee, J.-H. and Weinbach, G. C. (1993), Regime switching with time-varying transition probabilities, Working Papers 93-12, Federal Reserve Bank of Philadelphia.

Durland, J. M. and McCurdy, T. (1994), "Duration-dependent transitions in a markov model of u.s. gnp growth", *Journal of Business and Economic Statistics* , Vol. 12, pp. 279–88.

Durlauf, S. N. and Johnson, P. A. (1995), "Multiple regimes and cross-country growth behaviour", *Journal of Applied Econometrics* , Vol. 10, pp. 365–84.

Easterly, W., Kremer, M., Pritchett, L. and Summers, L. H. (1993), "Good policy or good luck?: Country growth performance and temporary shocks", *Journal of Monetary Economics* , Vol. 32, pp. 459–483.

Filardo, A. J. (1994), "Business-cycle phases and their transitional dynamics", *Journal of Business and Economic Statistics* , Vol. 12, pp. 299–308.

Hall, R. E. and Jones, C. I. (1999), "Why do some countries produce so much more output per worker than others?", *The Quarterly Journal of Economics* , Vol. 114, pp. 83–116.

Hamilton, J. D. (1990), "Analysis of time series subject to changes in regime", *Journal of Econometrics* , Vol. 45, pp. 39 – 70.

Hamilton, J. D. (1994), *Time Series Analysis*, 1 edn, Princeton University Press.

Hansen, B. E. (2000), "Sample splitting and threshold estimation", *Econometrica* , Vol. 68, pp. 575–604.

Hausmann, R., Pritchett, L. and Rodrik, D. (2004), Growth accelerations, Working paper series, Harvard University, John F. Kennedy School of Government.

Jerzmanowski, M. (2006), "Empirics of hills, plateaus, mountains and plains: A markov-switching approach to growth", *Journal of Development Economics* , Vol. 81, pp. 357–385.

Jones, B. F. and Olken, B. A. (2008), "The anatomy of start-stop growth", *The Review of Economics and Statistics* , Vol. 90, pp. 582–587.

Jong-A-Pin, R. and de Haan, J. (2007), Political regime change, economic reform and growth accelerations, Technical report.

Kerekes, M. (2009), "Growth miracles and failures in a markov switching classification model of growth", *Freie Universitat Berlin, Discussion Papers* , Vol. 11.

Knack, S. and Keefer, P. (1995), "Institutions and economic performance: Cross-country tests using alternative institutional measures", *Economics and Politics* , Vol. 7, pp. 207–227.

Kourtellos, A. (2002), A projection pursuit approach to cross country growth data, University of Cyprus Working Papers in Economics 0213, University of Cyprus Department of Economics.

Loh, W.-Y. (2002), Regression trees with unbiased variable selection and interaction detection, Technical report, Statistica Sinica.

Masanjala, W. H. and Papageorgiou, C. (2004), "The solow model with ces technology: nonlinearities and parameter heterogeneity", *Journal of Applied Econometrics* , Vol. 19, pp. 171–201.

Paap, R., Franses, P. H. and van Dijk, D. (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* , Vol. 77, pp. 553–570.

Pritchett, L. (2000), "Understanding patterns of economic growth: Searching for hills among plateaus, mountains, and plains", *World Bank Economic Review* , Vol. 14, pp. 221–50.

Pritchett, L. (2003), A toy collection, a socialist star, and a democratic dud? growth theory, vietnam, and the philippines., *in* '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* , Vol. 49, pp. 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* , Vol. 25, pp. 1100–1127.

Wacziarg, R. and Welch, K. H. (2003), Trade liberalization and growth: New evidence, Research papers, Stanford University, Graduate School of Business.

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