# Essays on Structural Change, Economic Takeoffs, and Growth Volatility

**BY**

**IOANNA STYLIANOU**

**DISSERTATION**

**SUBMITTED IN PARTIAL FULFILLMENT OF THE**

**REQUIREMENTS FOR THE DEGREE OF**

**DOCTOR OF PHILOSOPHY IN ECONOMICS**

**DEPARTMENT OF ECONOMICS**

**UNIVERSITY OF CYPRUS**

**MAY 2012**

# ΣΕΛΙΔΑ ΕΓΚΥΡΟΤΗΤΑΣ

**Υποψήφια Διδάκτορας: Ιωάννα Στυλιανού**

**Τίτλος Διατριβής:** Essays on Structural Change, Economic Takeoffs, and Growth Volatility.

Η παρούσα Διδακτορική Διατριβή εκπονήθηκε στο πλαίσιο των σπουδών για απόκτηση Διδακτορικού Διπλώματος στο Τμήμα Οικονομικών και εγκρίθηκε στις 21 Μαΐου 2012 από τα μέλη της Εξεταστικής Επιτροπής.

**Εξεταστική Επιτροπή**

**Ερευνητικός Σύμβουλος:**

'Αντρος Κούρτελλος, Επίκουρος Καθηγητής, Πανεπιστήμιο Κύπρου.

**'Αλλα μέλη**

'Ελενα Αντρέου, Αναπληρωτής Καθηγήτρια, Πανεπιστήμιο Κύπρου (Πρόεδρος).

Λούης Χριστοφίδης, Καθηγητής, Πανεπιστήμιο Κύπρου.

Carl-Johan Dalgaard, Professor, University of Copenhagen, Copenhagen, Denmark.

Steven N. Durlauf, Professor, University of Wisconsin, Madison, USA.

# Examination Committee

1. Associate Professor Elena Andreou - Chairman of the Committee
2. Professor  Louis Christofides - Internal Member
3. Professor Carl-Johan Dalgaard - External Mamber (University of Copenhagen, Copenhagen,  Denmark)
4. Professor Steven N. Durlauf - External Member (University of Wisconsin, Madison, USA)
5. Assistant  Professor  Andros Kourtellos - Supervisor

# Abstract


The common theme of my dissertation is how structural changes at the macroeconomic level affect economic growth and growth volatility. In particular my thesis focuses on the following three topics: In Chapter Two we investigate the univariate and the multivariate properties of key macroeconomic variables of the economy of Cyprus using state of the art statistical techniques that allow for structural breaks. In Chapter Three, we examine empirically if initial land inequality is an important determinant of take-off delays and long-run economic performance, while in Chapter Four, we uncover growth volatility regimes and identify their robust determinants. Chapter One includes a brief summary of the results derived from the above topics.

It is undisputable the fact that the Cyprus economy, has been subject to a number of substantial internal and external shocks and thus, in Chapter Two we investigate the impact of these shocks on its main macroeconomic time series. We do so drawing upon the large econometric literature that has determined how best to consider exogenous and endogenous breaks in the context of unit root testing. We consider the short-run as well as the long-run relationship between variables in the presence of structural breaks. Our results indicate that the null hypothesis of a unit root cannot be rejected for any of the series examined. The structural breaks found coincide with important, known, policy changes as well as economic and political events. While the short-run inter-relationship among variables is greatly affected by these events, we do not discern long-run effects on these relationships.

Recent work in the growth literature has provided various explanations for transition delays and the great divergence. Chapter Three provides empirical support for one theory of transition delays: initial land inequality. Our analysis is designed to elucidate the channels via which land inequality can affect long-run economic performance. Using a new historical data set for land inequality (Frankema (2009)) we employ duration analysis to investigate whether higher levels of land inequality lead to longer delays in the extension of primary schooling. We then investigate whether such delays affect long-run economic performance via their effect on contemporaneous schooling. Our findings suggest that land inequality is a key determinant of delays in schooling, and that such delays have a significant negative impact on long-run output.

In Chapter Four we uncover growth volatility regimes and identify their robust determinants using a large international panel of countries. In doing so we propose a novel empirical methodology that allows us to simultaneously deal with model uncertainty, heterogeneity, and endogeneity by unifying two recent econometric techniques: Bayesian Model Averaging and Structural Threshold Regression (STR). We find robust evidence for multiple volatility regimes indexed by levels of initial income and public debt. We also find heterogeneous but negative effects of institutions and demography on volatility across regimes.

Το βασικό θέμα αυτής της διατριβής είναι πως οι δομικές αλλαγές στο μακροοικονομικό επίπεδο επηρεάζουν την οικονομική ανάπτυξη, όπως επίσης, και τη διακύμανση των ρυθμών οικονομικής ανάπτυξης. Συγκεκριμένα, σε αυτή τη διατριβή εξετάζονται τρία θέματα: Στο Κεφάλαιο Δύο στα πλαίσια μιας μονομεταβλητής και πολυμεταβλητής ανάλυσης εξετάζονται οι ιδιότητες βασικών μακροοικονομικών μεταβλητών της Κυπριακής Οικονομίας εφαρμόζοντας στατιστικούς ελέγχους που επιτρέπουν την παρουσία ενδογενών και εξωγενών δομικών αλλαγών. Στο Κεφάλαιο Τρία, εξετάζουμε εμπειρικά εάν η κατανομή της γής αποτελεί καθοριστικό παράγοντα της αειφόρου και μακροχρόνιας οικονομικής ανάπτυξης, ενώ στο Κεφάλαιο Τέσσερα, αναλύουμε εάν οι χώρες μπορούν να κατανεμηθούν σε ομάδες με βάση τη διακύμανση των ρυθμών οικονομικής ανάπτυξης, ενώ ταυτόχρονα εξετάζουμε τους προσδιοριστικούς της παράγοντες. Το Κεφάλαιο Ένα περιλαμβάνει μια συνοπτική ανασκόπηση των αποτελεσμάτων που προέκυψαν από αυτή την ανάλυση.

Είναι αναμφισβήτητο το γεγονός ότι η Κυπριακή Οικονομία έχει υποστεί μια πλειάδα εσωτερικών και εξωτερικών σοκ και συνεπώς, στο Κεφάλαιο Δύο, εξετάζουμε την επίδραση αυτών των σοκ σε βασικές μακροοικονομικές μεταβλητές, λαμβάνοντας υπόψη την ύπαρξη τόσο ενδογενών όσο και εξωγενών δομικών αλλαγών σε ελέγχους μοναδιαίας ρίζας, όπως αυτοί προτείνονται από την βιβλιογραφία. Εξετάζουμε παράλληλα, την βραχυχρόνια και τη μακροχρόνια σχέση των μεταβλητών λαμβάνοντας υπόψη την παρουσία δομικών αλλαγών. Σύμφωνα με τα αποτελέσματα, η μηδενική υπόθεση της μοναδιαίας ρίζας δεν απορρίπτεται για όλες τις μεταβλητές, ενώ οι δομικές αλλαγές που εντοπίζονται, συνδέονται με σημαντικές αλλαγές νομισματικής πολιτικής, όπως επίσης και με γεγονότα οικονομικής και πολιτικής φύσεως. Οι δομικές αυτές αλλαγές επηρεάζουν τη βραχυχρόνια σχέση, αλλά όχι τη μακροχρόνια σχέση των μεταβλητών.

Πρόσφατες μελέτες στην βιβλιογραφία έχουν αναπτύξει διάφορες θεωρίες όσον αναφορά τις αιτίες που οι χώρες δεν μπορούν να επιτύχουν ένα επίπεδο αειφόρου ανάπτυξης και γενικά για την αποτυχία οικονομικής σύγκλισης μεταξύ των χωρών. Το Κεφάλαιο Τρία, στηρίζει εμπειρικά μια από αυτές τις θεωρίες: Την (αρχική) ανισότητα όσον αναφορά την κατανομή της γης. Στα πλαίσια αυτά, περιγράφουμε τις οδούς μέσα από τις οποίες η ανισότητα όσον αναφορά την κατανομή της γης επηρεάζει την μακροχρόνια οικονομική ανάπτυξη. Χρησιμοποιώντας ιστορικά στοιχεία για την κατανομή της γής (Frankema (2009)) και εφαρμόζοντας την "'Ανάλυση Διάρκειας Ζωής", εξετάζουμε εάν υψηλότερα επίπεδα ανισότητας στη γη, αποτελεί τροχοπέδη στην ανάπτυξη του επιπέδου της πρωτοβάθμιας εκπαίδευσης. Παράλληλα, εξετάζουμε εάν η επιβράδυνση αυτή επηρεάζει την μακροχρόνια οικονομική ανάπτυξη μέσω του επιπέδου εκπαίδευσης σήμερα. Σύμφωνα με τα ευρήματα μας, τα επίπεδα ανισότητας της γής επηρεάζουν αρνητικά τα επίπεδα εκπαίδευσης σήμερα καθώς επίσης, και το μακροχρόνιο προϊόν.

Στο Κεφάλαιο Τέσσερα αναλύουμε εάν οι χώρες μπορούν να κατανεμηθούν σε ομάδες με βάση τη διακύμανση των ρυθμών οικονομικής ανάπτυξης, ενώ ταυτόχρονα εξετάζουμε τους προσδιοριστικούς της παράγοντες χρησιμοποιώντας ένα μεγάλο αριθμό χωρών. Ως εκ τούτου, χρησιμοποιούμε μια μεθοδολογία, την Bayesian Model Averaging and Structural Threshold Regression (STR), που λαμβάνει υπόψη την αβεβαιότητα ως τον προσδιορισμό του μοντέλου (ετερογένεια και ενδογένεια). Σύμφωνα με τα αποτελέσματα, οι χώρες μπορούν να κατανεμηθούν ως προς τη διακύμανση των ρυθμών οικονομικής ανάπτυξης αν λάβουμε υπόψη τα επίπεδα του αρχικού κατακεφαλή ΑΕΠ και του δημόσιου χρέους. Ταυτόχρονα, τα επίπεδα των πολιτικών ιδρυμάτων και της δημογραφίας, επηρεάζουν αρνητικά τη διακύμανση των ρυθμών οικονομικής ανάπτυξης.

# Acknowledgments

Pursuing a Ph.D. is like climbing a high peak, step by step, accompanied with hardships, frustration, encouragement and trust and with so many people's kind help. When I found myself at the top I realized that it was, in fact, teamwork that got me there. Though it will not be enough to express my gratitude in words to all those people who helped me, I would still like to give my many, many thanks to all these people.

First of all, I would like to express my deepest sense of gratitude and appreciation to my mentor Professor Andros Kourtellos, for his guidance, patience and encouragement throughout the development of the thesis. I have been extremely lucky to have him as a supervisor and without his support and enthusiasm I would never be able to make this dream come true. I will be forever grateful and appreciative for everything that he has done for me.

I gratefully acknowledge Professor Louis Christofides for his advice, supervision, support, and crucial contribution, which made him a backbone of this research and so to this thesis. Further appreciation goes to Professor Elena Andreou for her invaluable support, guidance and understanding throughout my PhD.

I would also like to record my gratitude to Professor Steven N. Durlauf, for his supervision, advice, and guidance during my two-year visit at the University of Wisconsin-Maddison. Professor Durlauf, provided me unflinching encouragement and support in various ways. His truly scientist intuition has made him as a constant oasis of ideas and passions in science, which exceptionally inspire and enrich my growth as a student and as a researcher. I am indebted to him more than he knows.

I'd like to convey my heartfelt thanks to Professor Chih Ming Tan for his valuable advice in science discussion, help and support. I've learned a lot from him and I consider myself extremely lucky working with him.

Special thanks to the Department of Economics and its personnel. Especially, I want to express my gratitude to Professor Michael Michael, Costas Hadjiyiannis, and Sofronis Clerides, who kindly offered me their support and advice since 1999, creating the ideal environment to focus on my PhD.

I would also acknowledge my colleagues at the Department of Economics, Adamos Adamou, Neophyta Embora, Sofia Andreou, Kikis Agisilaou and Pantelis Mitsis for their advice and their willingness to share their bright thoughts with me.

There are two people I need to mention especially, Litsia Taliadorou and Anastasia Demetriou. First, for their precious friendship, but also, for their unselfish help and support all these years. I love them both so much and iam grateful to them in every possible way.

Many special thanks go to Andreas Andreou, Jozef Kontos and Maria Kanari, who generously offered me their help, especially in technical issues.

I also want to express my sincere gratitude to my dearest friend Maria. Her encouragement, love and loyalty gave me strength all these years and I owe her a lot. I really feel very lucky and proud having her as my friend.

Last but not least, I am greatly indebted to my family. They form the backbone and origin of my happiness and strength. Their love and support without any complaint or regret has enabled me to complete this PhD project. I am very grateful to my mother and I dedicate this thesis to her with all my love. Her understanding and her love encouraged me to work hard and to continue pursuing a Ph.D project here and in Madison. Her kind-hearted personality has affected me to be steadfast and never bend to difficulty. Despite the fact she is not here with us anymore, I know she would be very proud for me and happy for completing my PhD. Especially, I want to thank from the bottom of my heart, my sister Thekla, who took every responsibility and suffered all the bitterness in order to enable me to concentrate on my research. I love her so much and I owe her every achievement I made. Also, special thanks to my dearest uncle Michael and his wife Stella, to my aunt Agatha and her husband Andreas, and to my beloved cousins Ioanna and Katerina, who offered me their unconditional support and love in extremely difficult times.

*Dedication*

In loving memory of my mother Damiani Kleanthous

# Table of Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

In Chapter Two, we analyze the univariate and the multivariate properties of key macroeconomic variables of the economy of Cyprus over the period 1970q1-2010q1. First, we examine the unit root hypothesis under exogenous and under endogenous breaks and afterwards, we investigate not only the short-run, but also the long-run relationship between the variables, taking into account the presence of structural breaks. In particular, we consider multiple or single equation models that allow for endogenous breaks as in Bai and Perron (1998) and Qu and Perron (2007). These breaks are related to policy changes during the period examined, as well as to political and economic events in and out of Cyprus. Our results suggest that all variables are characterized by the existence of a unit root. Looking at the timing of the breaks, the global financial crisis, the interest rate liberalization, the Gulf War and various oil crises, have the most profound effect on the variables. In addition, the breaks affect significantly the short-run but not the long-run relationship of the variables.

Chapter Three focuses on the empirical support for one theory of transition delays , which is initial land inequality, proposed by Galor, Moav, and Vollrath (2009). Using a new historical data set for land inequality by Frankema (2009) we investigate if land inequality along with other candidates are important determinants on delays in primary schooling in the context of duration analysis and considering model uncertainty. In addition, we explore the question of how delays in schooling affect contemporaneous measures of schooling, institutions and long-run income via their effect on current institutions and schooling. Consistent with the theory, we find that increased levels of land inequality lead to more delays in reaching a specific primary schooling enrollment rate. Also, we find strong evidence that historical delays in achieving particular educational thresholds determine not only current schooling, but also current executive constraints. Finally, the results suggest that the historical evolutionary path of human capital is a fundamental determinant of long-run economic outcomes.

The objective of Chapter Four is to identify robust determinants of growth volatility using an econometric methodology which considers parameter heterogeneity and theory uncertainty by synthesizing model averaging and structural threshold regression (STR), recently proposed by Kourtellos, Stengos, and Tan (2011). According to the results, we find strong evidence for a threshold under various model specifications and using different endogenous variables, especially for Initial Income and Public Debt. In the STR estimation, initial life expectancy and institutions are the two key variables which affect negatively growth volatility, especially in high Income, high Inflation Volatility, and high Debt countries. The negative effect of institutions on Growth Volatility is also identified for low Inflation Volatility countries and in countries with low Public Debt, the impact of Life Expectancy is also strong. Finally, in countries with greater Trade Openness the effect of Life Expectancy and Population Growth is negative and significant whereas the negative effect of institutions on Growth Volatility is identified in countries with trade barriers.

# Chapter 2

# Structural Breaks in Applied Macroeconomic Modelling: The Case of Cyprus

## 2.1 Introduction

The main aim of this paper is to analyze the univariate and the multivariate properties of key macroeconomic variables of the economy of Cyprus, using state of the art statistical techniques that allow for structural breaks. Economic growth in Cyprus after 1960 has been remarkable and, according to Durlauf, Johnson and Temple (2005), this country is one of the fifteen growth miracles of the period 1960-2000.

Cyprus has a market economy dominated by the services sector, which accounts for four-fifths of GDP. Tourism, financial services and real estate are the most important sectors. Erratic growth rates over the past decade reflect the economy's reliance on tourism, the profitability of which often fluctuates with political instability in the region and economic conditions in Western Europe. Thus, Cyprus has a moderately high vulnerability to potential shocks. Furthermore, Cyprus has experienced wars and several institutional changes before and after its accession to the European Union. The invasion of Turkey, the Iraq and the Gulf War had a major impact on the Cyprus Economy. Also, Cyprus joined the European Exchange Rate Mechanism (ERM2) in May 2005 and adopted the euro as its national currency on 1 January 2008 whereas an aggressive austerity program in the preceding years, aimed at paving the way for the euro. Other events, like the global financial crisis had also a profound impact on Cyprus since construction and tourism slowed in the face of reduced foreign demand. All these events make Cyprus a particularly interesting case to study.

In this paper, we first examine the unit root hypothesis for Gross Domestic Product, two definitions of Money, the Deposit Rate, the Consumer Price Index, Employment, the Unemployment Rate, and the Exchange Rate. Variables which are characterized by a unit root suffer permanent effects from random shocks and have no tendency to return to a long-run deterministic path. We conduct unit root tests allowing for data-driven (endogenous) breaks as proposed by Zivot and Andrews (1992), Vogelsang and Perron (1998) and Kim and Perron (2009), or under exogenous breaks, following Perron (1989). We analyze the relationship between those variables not only in the short-run but also in the long-run, taking into account the presence of structural breaks. In particular, we consider multiple or single equation models that allow for endogenous breaks as in Bai and Perron (1998) and Qu and Perron

(2007). These breaks are related to policy changes during the period examined, as well as to political and economic events in and out of Cyprus.

One contribution of this study is to uncover the effect of these events on the main macroeconomic series of the Cyprus economy over the period 1970q1-2010q1. A second contribution is to investigate the unit root hypothesis under exogenous and endogenous break models. We also uncover the short-run and the long-run relationship between the variables studied using techniques which consider the presence of structural breaks.

The findings from this part are extremely significant in terms of economic policy implementation. Questions such as the following are addressed and the answers to these can offer useful input to policy makers: (i) Does economic policy affect important macroeconomic series and does it do so in a theoretically coherent way? (ii) Does this effectiveness, if present, differ before and after breaks that might be identified? (iii) What are the implications of breaks for the short-run and the long-run equilibrium relationship that may exist among variables?

Our results suggest that all variables are characterized by the existence of a unit root. Looking at the timing of the breaks, the global financial crisis, the interest rate liberalization, the Gulf War and various oil crises, have the most profound effect on the variables. In addition, the breaks affect significantly the short-run but not the long-run relationship of the variables.

This paper is organized as follows: Section 2.2 considers the main variables of the analysis and Section 2.3 develops the associated methodology and literature. Section 2.4 presents the unit root analysis under exogenous and endogenous breaks, as well as the multivariate analysis. Section 2.5 presents the conclusions derived from this study.



## 2.2 Data

We study eight quarterly macroeconomic series over the period 1970q1-2010q1. More precisely, we study the Gross Domestic Product (GDP), Money Liquidity (M2 and M1), the Deposit Rate, the Consumer Price Index (CPI), the Employed Population, the Unemployment Rate, and the Exchange Rate of the CYP/USD transformed into the Euro/USD rate using the standard relationship between the Euro and the CYP (0.585274) that prevailed during our period of study.

Table 2.1 describes the variables and their source, while Table 2.2 presents the descriptive statistics. Evidence for the presence of structural breaks is provided in Figures 2.1.1-2.1.8, where the variables are presented in first differences. These breaks are related with the various political and economic events that are presented in Table 2.3.

Figure 2.1.1 presents the GDP variable and obviously, the series displays a positive break in 1997q3. A detailed examination of the GDP components showed that the positive break is related with an expansion of the wholesale and retail trade during that period. According to Figures 2.1.2 and 2.1.3, M1 and M2 are affected from big structural breaks in 1999-2000 that can be attributed to the Cyprus Stock Exchange crisis. In particular, in 1996 the Cyprus Stock Exchange was established and a number of mature and relatively new companies were admitted to it, leading to speculative behaviour and a bubble.

In addition, for the Average Deposit Rate, presented in Figure 2.1.4, there is a big positive break in 2004q2 which is possibly due to the increase in the interest rate induced by the Central Bank of Cyprus as a reaction to the rumour that the Cyprus pound would be devalued. The political uncertainty which prevailed in Cyprus in April 2004, just before and after the Annan Plan referendum together with the full capital account liberalisation that occurred upon accession were the main reasons for the outbreak of these rumours.

The structural break related with the CPI (presented in Figure 2.1.5) in 1974q1 is possibly linked with the oil crisis which started in 1973q4, when the members of the Organization of Petroleum Exporting Countries proclaimed an oil embargo in response to the decision by the United States to re-supply the Israeli military during the Yom Kippur War.

In the labour market, Employment (Figure 2.1.6) and the Unemployment Rate (Figure 2.1.7) are characterized by two significant breaks in 2005q1, and in 1993q2, respectively. The increase in Employment in 2005q1 is related with the fact that Cyprus became a full member of the European Union, allowing any European citizen to work in the island. The Exchange Rate, which is presented in Figure 2.1.8, experienced a break in 1985q3 which was caused by the linkage of the Cyprus Pound to a trade-weighted currency basket.

## 2.3 Methodology

In this paper, we aim to uncover not only the univariate properties of our variables, but also their short-run as well as their long-run inter-relationships via the estimation of Vector Autoregressive (VAR) and Vector Error Correction Models (VECMs). However, it is significant to examine first if these variables are stationary, since the existence of unit roots has significant implications. The presence or absence of unit roots helps to identify some features of the underlying data-generating process. If the variables have unit roots, they are characterized as non-stationary processes that have no tendency to return to a long-run deterministic path and they have a time-dependent variance. In addition, non-stationary series suffer permanent effects from random shocks. Thus, any analysis which involves forecasting could be erroneous.

The standard procedure for detecting non-stationary behaviour is to use the Augmented Dickey-Fuller (ADF) test (Dickey and Fuller, 1979, 1981) using the following equation (assuming that we carry out the test for the GDP variable):

$$\Delta GDP_t = \mu + \beta t + \alpha GDP_{t-1} + \sum_{i=1}^{k} c_i \, \Delta GDP_{t-i} + \varepsilon_t \qquad (2.3.1)$$

Rejection of the null hypothesis ($a = 0$) implies that $GDP_t$ is stationary and integrated of order zero, or $I(0)$. This test, however, is likely to be biased towards not rejecting the unit root hypothesis as it does not account for the structural breaks that we have already observed in the graphical analysis. To allow for structural breaks, we first consider the breaks as exogenous (or known), and afterwards, as endogenous.

To test for a unit root under exogenous breaks we follow the methodology proposed by Perron (1989) who assumes that the dating of the break points is known a priori. Perron's (1989) procedure is characterized by a single exogenous (known) break in accordance with the underlying asymptotic distribution theory. Perron uses a modified Dickey-Fuller (DF) unit root tests that includes dummy variables to account for one known, or exogenous structural break. The break point of the trend function is fixed (exogenous) and chosen independently of the data.  Specifically, Perron suggests specifying the location of the break date, estimating a regression that nests the random walk null and the trend stationary alternative with either a change in intercept (Model A), a break in slope (Model B), or both (Model C).

In addition, we consider the unit root hypothesis when the break points are endogenously determined. We view the two approaches as complementary since the exogenous breaks approach is based on the prior information of the existence of the breaks while the endogenous breaks methodology allows the data select the breaks. Under the endogenous breaks scheme, Perron's methodology is applied for each possible break date in the sample, yielding a sequence of t-statistics. From this sequence, various algorithms can be used to construct 'minimum t-statistics' which maximize evidence against the null hypothesis. A methodology in line with this concept was proposed by Zivot and Andrews (1992) who considered the presence of a unit root in the intercept (Model A), the trend (Model B), or both (Model C). However, it is assumed that, if a break occurs, it does only under the alternative hypothesis of stationarity. As Kim and Perron (2009) pointed out, this is undesirable for several reasons. First, when allowing for a break, it imposes an asymmetric treatment, so that the test may reject when the noise is integrated but the trend is changing. Second, if a break is present, this information is not utilized to improve the power of the test. Kim and Perron (2009) address both issues by developing a unit root test which allows a break under the null and the alternative hypothesis. If a break is present, the limit distribution of the test is the same as in the case of a known break date, thereby allowing increased power and maintaining the correct size. They test the unit root hypothesis considering at length the notion of the Additive Outlier (AO), which applies to cases where the break is assumed to occur instantly and is not affected by the dynamics of the series, and the Innovational Outlier (IO) Model, which is applicable to cases where it is more reasonable to view the break as occurring more slowly over time. A Level Shift Model (Model A) and a Mixed Model (Model C) is available for both cases, whereas a Changing Growth Model (Model B) is applicable only for the Additive Outlier.

A significant part of the Kim and Perron (2009) unit root test is the identification of the presence of a break. Therefore, following Perron and Yabu (2009), we test the null hypothesis of a stable trend function, considering a structural change in the intercept (Model A), a change in the slope (Model B), or a change in both, intercept and trend (Model C), without having any knowledge about the stationarity of the noise component. If the null hypothesis is not rejected, the standard Augmented Dickey-Fuller (ADF) is utilized. Alternatively, if a break is present, the Kim and Perron (2009) unit root test is employed.

The findings regarding stationarity and the presence of structural breaks from the univariate analysis are considered in the second part of the paper where we analyze the dynamic inter-relationships between the variables. In particular, we estimate a VAR model which, since Sims' (1980) seminal paper, has been the standard tool in the analysis of the macroeconomic fluctuations. Considering the fact that the underlying variables are affected by structural breaks, we follow Qu and Perron (2007) who deal with the estimation of structural breaks in the mean, the covariance matrix of the errors, or both, in a system of equations. The estimation is carried out using a quasi-maximum-likelihood method, assuming serially uncorrelated Gaussian errors. In particular, conditional on a given partition of the sample $T = (T_{1,\dots\dots,}T_m)$, the Gaussian quasi-likelihood function is

$$L_T(T,\beta,\Sigma) = \prod_{j=1}^{m+1} \prod_{t=T_{j-1}+1}^{T_J} f\big(y_t/x_t;\ \beta_j, \Sigma_j\big) \qquad (2.3.2)$$

where

$$f\big(y_t/x_t; \beta_j, \Sigma_j\big) = \frac{1}{(2\pi)^{n/2}\left|\Sigma_j\right|^{1/2}} exp\left\{-\frac{1}{2}\left[y_t - x'_t\beta_j\right]'\Sigma_j^{-1}\left[y_t - x'_t\beta_j\right]\right\},$$

$\Sigma$ is the covariance matrix, and $m$ is the total number of structural changes in the system. Using a likelihood ratio test we examine the null hypothesis of no structural changes in the VAR, versus some specific number of changes. A significant assumption of this method is that the endogenous variables should be stationary. The estimated breaks will not only allow us to uncover the influence of various political and economic events on the dynamic inter-relationships between the variables but, also, to perform Impulse Response Function analysis, before and after the break, for economic policy evaluation purposes.

Finally, we examine the long-run equilibrium relationship between the variables via the estimation of a VECM. Therefore, we first test for the presence of co-integration and, as in the previous section, we consider the presence of structural breaks. If breaks are present, the standard co-integration tests are non-informative. Therefore, following Qu (2007), we test the null hypothesis of one stable co-integrating vector for the full sample against the alternative hypothesis of more than one co-integration relationships in some sub-sample. Qu (2007), proposes two types of tests: the WQ and the SQ tests if the number of the breaks are unknown, and the SupQm test, which allows for m changes.

## 2.4 Results

### 2.4.1 Unit Root Analysis Under Structural Breaks

We start our analysis by first applying the ADF test with a trend to the levels as well as to the first differences of the variables. Table 2.4 presents the results. Using the critical values tabulated by Dickey and Fuller (1979, 1981), we cannot reject the null hypothesis of a unit root for all variables for the level case. In addition, performing the test on the first differences, the unit root null is rejected for all variables, except for the M2 case. However, because of the weakness of the ADF unit root test to reject the unit root hypothesis in the presence of structural breaks, we investigate further the stationarity properties of the variables considering unit root tests under structural breaks.

Therefore, following Perron (1989) we assume the presence of an exogenous break and particularly, the Kosovo War in 1999q1. We could also set the exogenous break to be the Turkish invasion in 1974q3 or the Gulf War in 1991q1. However, for some variables the data started after 1974 or even after 1991. Thus, in order to use a common exogenous break for all variables we use the Kosovo War. The unit root results presented in Table 2.5, suggest that for all the variables the unit root hypothesis cannot be rejected for every model specification.

Also, we consider the breaks as endogenous, and employ the methodology proposed by Zivot and Andrews (1992) and Kim and Perron (2009). The endogenous breaks methodology is strictly related with the notion of the Additive Outlier (AO) and the Innovational Outlier (IO) Model, based on the time span of the break.

First, we consider the three alternative models proposed by Zivot and Andrews (1992), in the context of the IO model. According to the results (Table 2.6), the unit root hypothesis cannot be rejected for all variables, except for the case of Employment. The related min t-statistics for Model C, are presented in Figure 2.2, where noticeable is the fact that the variables are affected by multiple breaks. Considering the timing of the breaks, definitely the recent global financial crisis, the Cyprus Stock Exchange Crisis, the Gulf and the Iraq War, the liberalization of the interest rates in 2001, and the oil crisis, have the most profound effect on the variables. Our variables are characterized by different dynamics and thus, they are affected by different breaks.

In contrast to Zivot and Andrews (1992), Kim and Perron (2009) consider both the IO and the AO framework. In addition, they allow the presence of breaks not only under the null, but also, under the alternative hypothesis, increasing thereby the power of the test and maintaining the correct size and symmetry. An important step before the Kim and Perron (2009) analysis, is to identify the presence of breaks. Consequently, following Perron and Yabu (2009), we test the null hypothesis of a stable trend function, considering a structural change in the intercept (Model A), a change in the slope (Model B), or a change in both intercept and trend (Model C). The results are presented in Table 2.7, using both the Akaike and the Bayesian Information Criteria. According to the results, all variables are affected by breaks, except from two variables: The Unemployment Rate and the Exchange Rate. Therefore, for these two variables, we have to consider the results from the standard ADF test which showed the presence of a unit root. The most important breaks include the global financial crisis, the Cyprus Stock Exchange crisis, the liberalization of the interest rates, the new Monetary Policy framework introduced at the beginning of 1996, the Iraq War, the oil crisis, the entry of Cyprus in the European Union, and the political crisis in Derinia.

Table 2.8 presents the unit root test results of Kim and Perron (2009), using the Akaike and the Bayesian Information Criteria. In at least one case, the unit root null cannot be rejected for all variables. The events which had a major impact on the variables include the entrance of Cyprus into the Eurozone, the political crisis in Derinia, the Iraq and the Gulf War, the Stock Exchange crisis, and finally, the liberalization of the interest rates. Thus, some variables are driven by exogenous breaks, like the Gulf War, whereas others are affected from breaks related to economic changes.

An interesting question arising from the above analysis is: What are the properties in terms of stationarity of the first differences of the variables, since rejection of the unit root null implies that the series are first-order integrated? To answer this question, we employ the Kim and Perron (2009) unit root test on the first differences of the variables. Again, we perform the

Perron and Yabu (2009) structural break test on the first differences, presented in Table 2.9. Based on the findings, all the variables are affected from breaks, except the Exchange Rate. Considering though the ADF test, the unit root null is rejected. Table 2.10 presents the unit root test on the first differences following Kim and Perron (2009). In all cases, the results indicate that the unit root null is rejected. The importance of the findings regarding the stationarity becomes obvious in the next part, where we analyze the short and the long-run relationship between the variables.

### 2.4.2 Multivariate Analysis Under Structural Breaks

In this section, we examine first the short-run relationship among the variables, through the estimation of alternative VAR models considering the presence of structural breaks in the conditional mean, variance, or both, following the methodology proposed by Qu and Perron (2007).

The results for six alternative VAR models (with different money supply and labour force variables) are presented in Table 2.11. As Qu and Perron (2007) indicate, the variables should be stationary, and therefore, we employ the methodology using the first differences and allowing only one break due to the relatively small sample size. According to the results, the null hypothesis of no break versus one break is rejected for all cases. Interestingly, the Iraq War and the Cyprus Stock Exchange Crisis are the two most significant breaks affecting the models.

But the crucial question at this point is: What are the effects of the breaks in terms of economic policy? To deal with this issue, we split the sample before and after the break date and perform Impulse Response Function analysis. In what follows, using the model that includes M2 and Employment, we concentrate on monetary policy shocks which have particular policy interest looking at the period as a whole. According to Figure 2.3, a random shock in M2 has, in general, positive effects on GDP, the CPI, and Employment, before and after the Cyprus Stock Exchange Crisis, as predicted by the theory. However, there are significant differences in terms of the size and the span of the response before and after the break. Figures 2.3.1 and 2.3.2 show the Impulse Response Function for GDP before, and after the break. It is noticeable that, before the break, the positive response is more intense, since the size and the span are bigger than in the post-break era. The same results hold for the case of CPI presented in Figures 2.3.3 and 2.3.4, whereas for Employment (Figures 2.3.5 and 2.3.6), the response before the break is generally positive, but after the break, is negative, then positive and with shorter span than in the pre-break period. Finally, for the Deposit Rate (Figures 2.3.7 and 2.3.8) the effect before the break is positive and then negative whereas, after the break, the effect is positive but notably smaller in terms of size and span. Consequently, the presence of structural breaks and, in particular, the Cyprus Stock Exchange Crisis alters the effectiveness of monetary policy at least in the short-run.

Based on these findings, it is important to ask whether structural breaks affect the long-run equilibrium relationship between the variables? Qu (2007) points out that structural breaks can actually change the co-integrating relationship in some parts of the sample and he proposes two types of tests for the null hypothesis of a stable co-integrating relationship for the full sample. According to the results presented in Table 2.12, the null hypothesis of one stable co-integrating vector cannot be rejected under both a known and an unknown number of breaks. Therefore, despite the fact that in the short-run structural breaks alter the effectiveness of economic policy, they do not have any long-run implications.

In the next and final part of our paper, considering the presence of one stable co-integrating relationship, we estimate a VEC model based again on the model which includes M2 and Employment, allowing a trend in the co-integrating equation as well as in the error correction term. According to Table 2.13, we obtain the following co-integrated equation:

$$GDP = 0.143 * M2 + 0.006 * DEPOSIT - 3.512 * CPI + 0.197 * EMPLOYMENT + 0.028 * t + 31.052 \quad (2.4.1)$$

The above equation summarizes the long-run relationship between the variables which is consistent with the theory: GDP is positively related with M2, the Deposit Rate, and Employment, but for the case of CPI the relationship is negative. CPI is related positively with M2, the Deposit Rate, and Employment, M2 is negatively linked with the Deposit Rate and Employment, and finally, Employment is influenced negatively by the Deposit Rate. Based on this model, we perform also a cumulative Impulse Response Function analysis (Figure 2.4). According to the results, a random innovation in M2 has positive effects on GDP, Employment and CPI. In addition, as theory predicts, a random innovation in M2 affects negatively the Deposit Rate.

## 2.5 Conclusion

The unstable environment inside and outside Cyprus, greatly influenced by political and economic events, appears to have affected key macroeconomic variables (GDP, M1, M2, the Deposit Rate, the CPI, Employment, the Unemployment Rate, and the Exchange Rate) of the Cyprus economy. We, therefore, consider these events and attempt to take their influence into account in our univariate and multivariate analysis.

In the first part of this study, where we examine the unit root hypothesis, we take into account the presence of endogenous and exogenous breaks. The results from testing the unit root hypothesis under exogenous and endogenous breaks do not provide much evidence against the null of a unit root. Furthermore, the unit root tests under endogenous breaks provide evidence for known events that could have been expected to have an effect on the Cyprus Economy. The recent global financial crisis, the liberalization of interest rates, the Cyprus stock exchange crisis, various oil crises, the Gulf and the Iraq war, and the entrance of Cyprus into the European Union, all appear to have affected significantly the behaviour of the macroeconomic variables that we study.

We then consider the inter-relationships among these variables in multivariate schemes and under structural breaks. In particular, we estimate alternative VAR models taking into account the presence of structural breaks in the conditional mean, variance, or both. According to the results, the Cyprus stock exchange crisis and the Iraq war play a crucial role in these models. An Impulse Response Function analysis showed that the response of the variables under monetary policy shocks differs greatly, in terms of size and span, before and after the break date. Consequently, the presence of structural breaks, at least in the short-run, affects significantly the effectiveness of economic policy. Furthermore, we examine if the breaks have any implications in the long-run by testing for the existence of a stable co-integrating relationship. In contrast with the short-run, the breaks did not affect the long-run equilibrium relationship.

In the last part of the paper and taking into account the co-integration findings, we estimate a VECM. Our results suggest that GDP is positively related with M2, the Deposit Rate, and Employment, but, in the case of the CPI, the relationship is negative. The CPI is related positively with M2, the Deposit Rate, and Employment. M2 is negatively linked with the Deposit Rate and Employment, and, finally, Employment is influenced negatively by the

Deposit Rate. An Impulse Response Function analysis shows that a random innovation in M2 has positive effects on GDP, Employment and the CPI, whereas the effect on the Deposit Rate is negative.

Our results represent a significant source of information on the influence of major events on the macroeconomy and offer helpful guidance on the formation of policy.

**Table 2.1: Variable Description and Sources**

| *Variable* | *Time Span* | *Description* | *Source* |
|---|---|---|---|
| Gross Domestic Product | 1995q1-2010q1 | Gross Domestic Product at 2005 constant prices (Chain Linking Method), Millions of Euros | Cyprus Statistical Service |
| Money Supply (M2) | 1980q1-2007q4 | Nominal Money Supply in Millions of Euros | Global Financial Data |
| Money Supply (M1) | 1980q1-2007q4 | Nominal Money Supply in Millions of Euros | Global Financial Data |
| Deposit Rate | 1970q4-2010q1 | Time Deposit Rate (3-Month) | Global Financial Data |
| Consumer Price Index (CPI) | 1970q1-2010q1 | Consumer Price Index | Global Financial Data |
| Employment | 1991q1-2009q4 | Thousands of Employed People | International Monetary Fund |
| Unemployment Rate | 1991q1-2009q4 | Ratio of the number of the Unemployed people to the total Labour Force | International Monetary Fund |
| Exchange Rate (EURO/USD) | 1970q1-2010q1 | Exchange rate of CYP to USD. The ratio was afterwards transformed to EURO/USD using the standard EURO=0.585274 CYP | Global Financial Data |

**Table 2.2: Descriptive Statistics**

| *Variable* | *Mean* | *St. Deviation* | *Minimum* | *Maximum* |
|---|---|---|---|---|
| Gross Domestic Product (in million Euros) | 3080 | 490 | 2220 | 3960 |
| Money Supply M2 (in million Euros ) | 7070 | 5890 | 803 | 23400 |
| Money Supply M1 (in million Euros) | 1190 | 980 | 217 | 4430 |
| Deposit Rate | 0.052 | 0.010 | 0.021 | 0.065 |
| Consumer Price Index (CPI) | 61.950 | 29.021 | 17.149 | 112.600 |
| Employment | 310581 | 43228 | 174500 | 385700 |
| Unemployment Rate | 0.032 | 0.006 | 0.017 | 0.049 |
| Exchange Rate (EURO/USD) | 0.794 | 0.133 | 0.578 | 1.163 |

Note: The data in this table are not seasonally adjusted. In the following analysis however, all variables are in logarithmic form and filtered for seasonality (except from the Deposit Rate where the seasonality null was rejected) using X12 additive ARIMA.

**Table 2.3: Chronology of Events**

| *Date* | *Description* |
|---|---|
| 1960-1972 | Cyprus Pound was pegged to Sterling |
| 1972-1973 | Cyprus Pound was pegged to USD |
| 1973:q4 | Oil Crisis |
| 1973-1984 | Cyprus Pound was linked to an import-weighted currency basket |
| 1974:q3 | Turkish Invasion in Cyprus |
| 1979-1980 | Oil Crisis |
| 1984-1992 | Cyprus Pound was linked to a trade-weighted currency basket |
| 1986:q1 | Oil Price Collapse |
| 1987:q4 | A Protocol for the second stage of the Association Agreement with the European Economic Community was signed in Luxembourg |
| 1990 | Oil Crisis |
| 1991:q1 | Gulf War |
| 1992:q2 | The peg of Cyprus Pound changed from a broad basket of currencies to the ECU |
| 1992:q3 | Pound Sterling and Italian Lira forced out of the European Exchange Rate Mechanism |
| 1996:q1 | The Central Bank put in to effect a new framework of Monetary Policy implementation |
| 1996:q3 | Political crisis due to the conflict in Derinia. |
| 1999:q1 | Kosovo War |
| 1999:q1 | The EURO replaced the ECU as the basis of the peg |
| 1999:q4 -2000:q1 | Cyprus Stock Exchange Crisis |
| 2000:q3-onwards | Cyprus Government put into effect a new framework in calculating oil price The price was liberalized in 2004:2 |

**Table 2.3-Continued**

| *Date* | *Description* |
|---|---|
| 2001:1 | Liberalization of the interest rates. Introduction of wider fluctuation margins for the Euro (from 2.25% to 15%) |
| 2001:q3-2001:q4 | US terrorism attack. As a consequence of the global economic slowdown following the 11 September attacks, interest rates were reduced in September and November 2001 by 50 basis points in each case |
| 2003:q1 | Iraq War |
| | Cyprus became a full member of the EU |
| | Annan Plan referendum |
| 2004:q2 | The political uncertainty which prevailed in Cyprus in April 2004 just before and after the Annan Plan referendum, together with the full capital account liberalisation upon accession, were the main reasons for the outbreak of rumours of a possible devaluation of the Cyprus pound. The Central Bank of Cyprus reacted to these rumours, which caused limited but persistent capital outflows, by comments by the Governor intended to send appropriate signals, as well as by increasing interest rates by 100 basis points |
| 2005:q2 | Cyprus entered ERM II |
| 2006:q3 | Increase in the interest by 25 basis points due to high inflationary pressures from the high oil prices and credit growth |
| 2007-onwards | Global Financial Crisis. However, Cyprus was affected greatly form the crisis during 2009. |
| 2008:q1 | Cyprus entered the Eurozone |

**Table 2.4: Dickey and Fuller (1979, 1981) Unit Root Test**

| *Variable* | Lags | $\hat{\mu}$ | $t_{\hat{\mu}}$ | $\hat{\beta}$ | $t_{\hat{\beta}}$ | $\hat{a}$ | $t_{\hat{a}}$ | *Critical Values* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | 1% | 5% | 10% |
| GDP | 2 | 1.125 | 0.57 | 0.000 | 0.46 | -0.053 | -0.56 | -4.132 | -3.492 | -3.175 |
| M2 | 5 | 0.641 | 2.03 | 0.000 | 1.95 | -0.032 | -1.98 | -4.038 | -3.449 | -3.149 |
| M1 | 2 | 1.412 | 1.62 | 0.001 | 1.63 | -0.076 | -1.60 | -4.037 | -3.449 | -3.149 |
| Deposit Rate | 4 | 0.056 | 0.61 | -0.000 | -2.29 | 0.014 | 0.45 | -4.022 | -3.443 | -3.143 |
| CPI | 8 | 0.059 | 2.03 | 0.000 | 0.54 | -0.013 | -1.46 | -4.023 | -3.443 | -3.143 |
| Employment | 1 | 1.549 | 2.17 | 0.000 | 2.01 | -0.129 | -2.15 | -4.097 | -3.476 | -3.166 |
| Unemployment Rate | 1 | -0.504 | -2.59 | 0.000 | 1.72 | -0.124 | -2.63 | -4.097 | -3.476 | -3.166 |
| Euro/USD | 3 | -0.019 | -1.72 | 0.000 | 0.85 | -0.056 | -2.44 | -4.021 | -3.442 | -3.142 |
| DGDP | 1 | 0.019 | 1.65 | -0.000 | -1.21 | -0.822 | -3.840 | -4.132 | -3.492 | -3.175 |
| DM2 | 4 | 0.015 | 1.47 | -0.000 | -0.30 | -0.455 | -2.05 | -4.038 | -3.449 | -3.149 |
| DM1 | 1 | 0.017 | 1.41 | 0.000 | 0.57 | -0.886 | -6.18 | -4.037 | -3.449 | -3.149 |
| DDeposit Rate | 3 | 0.015 | 1.90 | -0.000 | -2.87 | -1.175 | -8.10 | -4.022 | -3.443 | -3.143 |
| DCPI | 7 | 0.018 | 4.69 | -0.000 | -3.74 | -0.893 | -5.17 | -4.023 | -3.443 | -3.143 |
| DEmployment | 2 | -0.001 | -0.19 | 0.000 | 0.86 | -1.288 | -7.53 | -4.102 | -3.478 | -3.167 |
| DUnemployment Rate | 2 | 0.009 | 0.21 | -0.000 | -0.02 | -0.488 | -4.06 | -4.102 | -3.478 | -3.167 |
| DEuro/USD | 1 | 0.003 | 0.49 | -0.000 | -0.50 | -0.830 | -8.70 | -4.021 | -3.442 | -3.142 |

**Table 2.5: Perron (1989) Unit Root Test**

| *Variables* | T | λ | Lags | $\hat{\mu}$ | $t_{\hat{\mu}}$ | $\hat{\theta}$ | $t_{\hat{\theta}}$ | $\hat{\beta}$ | $t_{\hat{\beta}}$ | $\hat{d}$ | $t_{\hat{d}}$ | $\hat{a}$ | $t_{\hat{a}}$ | *Critical Values* | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | | | 1% | 5% | 10% |
| Model A | | | | | | | | | | | | | | | | |
| GDP | 61 | 0.278 | 4 | 7.592 | 3.172 | 0.009 | 2.523 | 0.003 | 2.928 | -0.006 | 1.963 | -0.366 | 3.160 | -4.39 | -3.76 | -3.46 |
| M2 | 112 | 0.687 | 1 | 0.820 | 1.690 | -0.004 | 0.745 | 0.001 | 1.521 | -0.000 | 0.064 | -0.041 | 1.640 | -4.42 | -3.80 | -3.51 |
| M1 | 112 | 0.687 | 1 | 0.930 | 1.074 | 0.022 | 1.321 | 0.000 | 0.878 | -0.024 | 1.336 | -0.048 | 1.004 | -4.42 | -3.80 | -3.51 |
| Deposit Rate | 158 | 0.721 | 4 | -0.047 | 0.354 | -0.038 | 1.636 | -0.000 | 0.130 | 0.038 | 1.539 | -0.018 | 0.329 | -4.42 | -3.80 | -3.51 |
| CPI | 161 | 0.726 | 4 | 0.048 | 1.451 | 0.000 | 0.120 | 0.000 | 0.264 | 0.002 | 0.626 | -0.010 | 0.956 | -4.42 | -3.80 | -3.51 |
| Employment | 76 | 0.434 | 1 | 1.458 | 1.813 | 0.003 | 0.440 | 0.000 | 1.279 | -0.007 | 1.548 | -0.121 | 1.782 | -4.34 | -3.72 | -3.44 |
| Unemployment | 76 | 0.434 | 1 | -0.594 | 2.775 | 0.035 | 1.622 | 0.001 | 2.205 | 0.091 | 5.604 | -0.128 | 2.404 | -4.34 | -3.72 | -3.44 |
| Exchange Rate | 161 | 0.726 | 1 | -0.018 | 1.286 | -0.013 | 1.175 | 0.000 | 1.178 | 0.041 | 4.623 | -0.044 | 1.508 | -4.42 | -3.80 | -3.51 |
| Model B | | | | | | | | | | | | | | | | |
| GDP | 61 | 0.278 | 4 | 5.442 | 2.278 | - | - | 0.002 | 2.715 | - | - | -0.266 | 2.311 | -4.51 | -3.87 | -3.58 |
| M2 | 112 | 0.687 | 1 | 1.299 | 1.878 | - | - | 0.002 | 1.722 | - | - | -0.066 | 1.833 | -4.51 | -3.85 | -3.57 |
| M1 | 112 | 0.687 | 1 | 2.339 | 2.146 | - | - | 0.002 | 1.948 | - | - | -0.124 | 2.106 | -4.51 | -3.85 | -3.57 |
| Deposit Rate | 158 | 0.721 | 4 | -0.553 | 1.776 | - | - | 0.000 | 1.137 | - | - | -0.188 | 1.793 | -4.51 | -3.85 | -3.57 |
| CPI | 161 | 0.726 | 4 | 0.053 | 1.415 | - | - | 0.000 | 0.410 | - | - | -0.012 | 0.983 | -4.51 | -3.85 | -3.57 |
| Employment | 76 | 0.434 | 1 | 3.575 | 3.235 | - | - | 0.000 | 2.403 | - | - | -0.291 | 3.225 | -4.55 | -3.94 | -3.66 |
| Unemployment | 76 | 0.434 | 1 | -0.692 | 1.522 | - | - | 0.001 | 0.836 | - | - | -0.151 | 1.889 | -4.55 | -3.94 | -3.66 |
| Exchange Rate | 161 | 0.726 | 1 | -0.036 | 2.235 | - | - | 0.000 | 2.384 | - | - | -0.077 | 2.326 | -4.51 | -3.85 | -3.57 |

**Table 2.5: Perron (1989) Unit Root Test-*Continued***

| *Variables* | T | λ | Lags | $\hat{\mu}$ | $t_{\hat{\mu}}$ | $\hat{\theta}$ | $t_{\hat{\theta}}$ | $\hat{\beta}$ | $t_{\hat{\beta}}$ | $\hat{d}$ | $t_{\hat{d}}$ | $\hat{a}$ | $t_{\hat{a}}$ | *Critical Values* | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | | | 1% | 5% | 10% |
| Model C | | | | | | | | | | | | | | | | |
| GDP | 61 | 0.278 | 4 | 8.078 | 2.935 | -0.033 | 0.458 | 0.002 | 2.720 | -0.006 | 2.006 | -0.388 | 2.943 | -4.78 | -4.17 | -3.87 |
| M2 | 112 | 0.687 | 1 | 1.412 | 1.939 | 0.041 | 0.593 | 0.002 | 1.818 | -0.003 | 0.479 | -0.072 | 1.899 | -4.75 | -4.18 | -3.86 |
| M1 | 112 | 0.687 | 1 | 2.446 | 1.782 | -0.286 | 1.181 | 0.002 | 1.583 | 0.003 | 0.120 | -0.130 | 1.745 | -4.75 | -4.18 | -3.86 |
| Deposit Rate | 158 | 0.721 | 4 | -0.532 | 1.655 | 0.455 | 1.576 | 0.000 | 1.533 | 0.028 | 1.128 | -0.180 | 1.661 | -4.75 | -4.18 | -3.86 |
| CPI | 161 | 0.726 | 4 | 0.050 | 1.212 | 0.003 | 0.181 | 0.000 | 0.269 | 0.018 | 0.516 | -0.011 | 0.810 | -4.75 | -4.18 | -3.86 |
| Employment | 76 | 0.434 | 1 | 3.485 | 2.792 | -0.143 | 2.285 | 0.000 | 1.719 | -0.003 | 0.609 | -0.284 | 2.763 | -4.81 | -4.22 | -3.95 |
| Unemployment | 76 | 0.434 | 1 | -0.854 | 1.759 | 0.145 | 0.521 | 0.003 | 1.272 | 0.087 | 5.371 | -0.164 | 2.002 | -4.81 | -4.22 | -3.95 |
| Exchange Rate | 161 | 0.726 | 1 | -0.038 | 2.202 | 0.174 | 1.731 | 0.000 | 2.251 | 0.017 | 1.032 | -0.085 | 2.324 | -4.75 | -4.18 | -3.86 |

Note: We regress using as TB the Kosovo War. The critical values are based on Perron´s (1989) Tables IV.B, V.B, and VI.B, at 1%, 5% and 10%. Model A involves a break in the intercept, Model B a break in the slope, whereas Model C a break in both intercept and slope.

**Table 2.6: Zivot and Andrews (1992) Unit Root Test**

| *Variable* | *Min t-statistic* | *Break* |
|---|---|---|
| *Model A* | | |
| GDP | -2.885 | 2009q2 |
| M2 | -3.872 | 2000q2 |
| M1 | -3.487 | 2003q3 |
| Deposit Rate | -3.235 | 2001q1 |
| CPI | -3.470 | 1978q4 |
| Employment | -23.656 | 2005q1 |
| Unemployment Rate | -3.47 | 1993q1 |
| Euro/USD | -3.210 | 1981q1 |
| *Model B* | | |
| GDP | -3.224 | 2008q4 |
| M2 | -2.970 | 1994q1 |
| M1 | -4.093 | 2002q4 |
| Deposit Rate | -3.538 | 1998q4 |
| CPI | -3.831 | 1983q2 |
| Employment | -19.022 | 2001q3 |
| Unemployment Rate | -2.988 | 1993q3 |
| Euro/USD | -2.676 | 2000q3 |
| *Model C* | | |
| GDP | -2.971 | 2007q4 |
| M2 | -3.123 | 1991q2 |
| M1 | -4.513 | 2000q4 |
| Deposit Rate | -4.711 | 2001q1 |
| CPI | -4.346 | 1979q3 |
| Employment | -23.918 | 2005q1 |
| Unemployment Rate | -3.578 | 1993q1 |
| Euro/USD | -3.137 | 1981q1 |

Note: The critical values for Model A are -5.43 and -4.80, for Model B -4.93 and -4.42, and for Model C, -5.57 and -5.08, for 1% and 5%, respectively. Model A involves a break in the intercept, Model B a break in the trend, whereas Model C a break in both intercept and trend.

**Table 2.7: Perron and Yabu (2009) Structural Break Test (Levels)**

| | *BIC* | | *AIC* | |
|---|---|---|---|---|
| | *Exp-Statistic* | *Break* | *Exp-Statistic* | *Break* |
| *Model A* | | | | |
| GDP | 0.1388 | 1999q2 | 2.5299** | 1999q2 |
| M2 | 1.4089* | 2003q1 | 3.2969*** | 2003q1 |
| M1 | 5.1267*** | 2003q3 | 5.3729*** | 2003q3 |
| Deposit Rate | 16.6350*** | 2001q2 | 17.1536*** | 2001q2 |
| CPI | 0.6012 | 1979q2 | 0.5688 | 1979q2 |
| Employment | 2.6545** | 2004q4 | 2.6545** | 2004q4 |
| Unemployment Rate | 0.0635 | 1995q4 | 0.4492 | 1995q4 |
| Euro/USD | 0.9317 | 1981q2 | 1.1037 | 1981q2 |
| *Model B* | | | | |
| GDP | 1.1097 | 2007q3 | 1.0058 | 2007q3 |
| M2 | 0.9246 | 1996q2 | 23.7067*** | 1996q2 |
| M1 | 15.3756*** | 2003q2 | 15.3756*** | 2003q2 |
| Deposit Rate | 3.5021*** | 1999q1 | 40.2673*** | 1999q1 |
| CPI | 10.3472*** | 1984q3 | 22.0156*** | 1984q3 |
| Employment | 1.2718* | 2003q1 | 1.1699* | 2003q1 |
| Unemployment Rate | -0.1147 | 1997q4 | -0.0992 | 1997q4 |
| Euro/USD | 0.2765 | 2001q4 | 0.1007 | 2001q4 |
| *Model C* | | | | |
| GDP | 1.9902 | 2007q3 | 6.7427*** | 2007q3 |
| M2 | 2.8375* | 1996q1 | 22.7549*** | 1996q1 |
| M1 | 16.1418*** | 2001q4 | 16.1418*** | 2001q4 |
| Deposit Rate | 21.7656*** | 2000q4 | 25.8068*** | 2000q4 |
| CPI | 10.8682*** | 1984q3 | 25.1066*** | 1984q3 |
| Employment | 4.2427** | 1996q3 | 4.2427** | 1996q3 |
| Unemployment Rate | 0.3185 | 2000q1 | 0.5294 | 2000q1 |
| Euro/USD | 1.9225 | 1999q4 | 1.2832 | 1999q4 |

Note: ***, ** and * denote rejection of the stable trend null at the 1%, 5% and 10% respectively. Critical Values for Model A: 1.260, 1.740, 3.120 for 10, 5, and 1% respectively. Critical Values for Model B: 1.130, 1.670 and 3.060 for 10, 5, and 1% respectively. Critical Values for Model C: 2.480, 3.120, 4.470 10, 5, and 1% respectively. Model A involves a break in the intercept, Model B a break in the slope, whereas Model C a break in both intercept and slope.

**Table 2.8: Perron and Kim (2009) Unit Root Test (Levels)**

| | *BIC* | | | *AIC* | | |
|---|---|---|---|---|---|---|
| | ***k*** | *Break (TB/T)* | *Unit Root Test* | *k* | *Break (TB/T)* | ***Unit Root Test*** |
| *Additive Outlier Changing Growth Model* | | | | | | |
| GDP | 0 | 2008q4(0.92) | -2.9295 | 0 | 2008q4(0.92) | -2.9295 |
| M2 | 1 | 1996q2(0.58) | -2.7250 | 10 | 1996q2(0.58) | -4.9028* |
| M1 | 3 | 2003q2(0.83) | -5.2635* | 3 | 2003q2(0.83) | -5.2635* |
| Deposit Rate | 0 | 1999q1(0.72) | -3.5103 | 10 | 1999q1(0.72) | -4.5844* |
| CPI | 4 | 1984q3(0.36) | -2.6749 | 8 | 1984q3(0.36) | -2.0774 |
| Employment | 0 | 1991q3(0.03) | -5.1533* | 0 | 1991q3(0.03) | -5.1533* |
| *Additive Outlier Mixed Model* | | | | | | |
| GDP | 0 | 2008q1(0.87) | -2.8867 | 0 | 2008q1(0.87) | -2.8867 |
| M2 | 3 | 1996q1(0.58) | -2.7039 | 10 | 1996q1(0.58) | -4.7735* |
| M1 | 0 | 2001q4(0.78) | -3.9902** | 7 | 2001q4(0.78) | -3.6654 |
| Deposit Rate | 1 | 2000q4(0.76) | -4.4245* | 9 | 2000q4(0.76) | -3.8260** |
| CPI | 4 | 1984q3(0.36) | -2.6898 | 8 | 1984q3(0.36) | -2.1014 |
| Employment | 0 | 1991q3(0.03) | -4.9883* | 0 | 1991q3(0.03) | -4.9883* |
| *Innovational Outlier Mixed Model* | | | | | | |
| GDP | 0 | 2006q4(0.79) | -2.8471 | 0 | 2006q4(0.79) | -2.8471 |
| M2 | 1 | 1997q1(0.61) | -2.4701 | 10 | 1997q1(0.61) | -4.9272* |
| M1 | 0 | 1999q3(0.70) | -3.9770** | 3 | 1999q3(0.70) | -5.1510* |
| Deposit Rate | 2 | 2000q4(0.76) | -4.5832* | 15 | 2000q4(0.76) | -4.4136* |
| CPI | 0 | 1985q3(0.39) | -2.1493 | 8 | 1985q3(0.39) | -1.8544 |
| Employment | 1 | 1991q3(0.03) | -2.0935 | 1 | 1991q3(0.03) | -2.0935 |

Note: * and ** denote a rejection of the unit root null at 5% and 10% respectively. k refers to the number of the lags of the differences of the dependent variable.

**Table 2.9: Perron and Yabu (2009) Structural Break Test (Differences)**

| | *BIC* | | *AIC* | |
|---|---|---|---|---|
| | *Exp-Statistic* | *Break* | *Exp-Statistic* | *Break* |
| *Model A* | | | | |
| GDP | 2.3052** | 2007q4 | 4.1884*** | 2007q4 |
| M2 | 0.0363 | 1994q3 | 5.9252*** | 1994q3 |
| M1 | 2.8385** | 2002q3 | 2.5261** | 2002q3 |
| Deposit Rate | 0.3673 | 2000q4 | 0.4210 | 2000q4 |
| CPI | 0.7549 | 1981q4 | 0.9974 | 1981q4 |
| Employment | 0.3277 | 2003q1 | 1.1104 | 2003q1 |
| Unemployment Rate | 0.2787 | 1999q3 | 0.2736 | 1999q3 |
| Euro/USD | 0.4493 | 2001q2 | 0.7506 | 2001q2 |
| *Model B* | | | | |
| GDP | 6.7989*** | 2007q4 | 6.7965*** | 2007q4 |
| M2 | 0.9835 | 2003q3 | 0.7979 | 2003q3 |
| M1 | 1.8217** | 2001q1 | 1.4290* | 2001q1 |
| Deposit Rate | 0.6301 | 2004q1 | 0.5684 | 2004q1 |
| CPI | -0.1388 | 1978q4 | -0.1308 | 1978q4 |
| Employment | 5.2471*** | 1994q2 | 8.7497*** | 1994q2 |
| Unemployment Rate | 1.1622* | 1994q2 | 1.1622* | 1994q2 |
| Euro/USD | 0.1209 | 1981q2 | 0.4999 | 1981q2 |
| *Model C* | | | | |
| GDP | 7.7045*** | 2006q4 | 7.7046*** | 2006q4 |
| M2 | 7.0709*** | 2002q4 | 7.2926*** | 2002q4 |
| M1 | 3.4354** | 2000q2 | 4.7527*** | 2000q2 |
| Deposit Rate | 1.6977 | 2004q1 | 4.2670** | 2004q1 |
| CPI | 2.6079* | 1981q4 | 3.6858** | 1981q4 |
| Employment | 8.4975*** | 1994q2 | 8.6535*** | 1994q2 |
| Unemployment Rate | 2.2822 | 1994q2 | 2.2597 | 1994q2 |
| Euro/USD | 1.5649 | 1985q1 | 1.6411 | 1985q1 |

Note: ***, ** and * denote rejection of the stable trend null at the 1%, 5% and 10% respectively. Critical Values for Model A: 1.260, 1.740, and 3.120 for 10, 5, and 1% respectively. Critical Values for Model B: 1.130, 1.670 and 3.060 for 10, 5, and 1% respectively. Critical Values for Model C: 2.480, 3.120, 4.470 10, 5, and 1% respectively. Model A involves a break in the intercept, Model B a break in the slope, whereas Model C a break in both intercept and slope.

**Table 2.10: Perron and Kim (2009) Unit Root Test (Differences)**

| | *BIC* | | | *AIC* | | |
|---|---|---|---|---|---|---|
| | *k* | *Break (TB/T)* | *Unit Root Test* | *k* | *Break (TB/T)* | *Unit Root Test* |
| *Additive Outlier Changing Growth Model* | | | | | | |
| GDP | 0 | 2007q4(0.85) | -11.1733* | 0 | 2007q4(0.85) | -11.1733* |
| M2 | 0 | 2005q1(0.90) | -8.1932* | 0 | 2005q1(0.90) | -8.1932* |
| M1 | 0 | 2001q1(0.76) | -12.4870* | 4 | 2001q1(0.76) | -5.6940* |
| Deposit Rate | 3 | 2007q2 (0.93) | -8.9423* | 3 | 2007q2 (0.93) | -8.9423* |
| CPI | 3 | 1973q3(0.09) | -5.9100* | 7 | 1973q3(0.09) | -5.0287* |
| Employment | 0 | 1991q4(0.04) | -13.5578* | 1 | 1991q4(0.04) | -11.0123* |
| Unemployment Rate | 0 | 1992q3(0.08) | -6.0135* | 0 | 1992q3(0.08) | -6.0135* |
| *Additive Outlier Mixed Model* | | | | | | |
| GDP | 0 | 2008q3(0.90) | -11.4796* | 0 | 2008q3(0.90) | -11.4796* |
| M2 | 0 | 2002q4(0.82) | -8.5697* | 3 | 2002q4(0.82) | -6.2700* |
| M1 | 0 | 2000q2(0.73) | -11.1536* | 4 | 2000q2(0.73) | -4.0842** |
| Deposit Rate | 3 | 2008q4(0.96) | -8.6620* | 3 | 2008q4(0.96) | -8.6620* |
| CPI | 3 | 1981q4(0.29) | -6.5924* | 7 | 1981q4(0.29) | -5.8820* |
| Employment | 0 | 1991q4(0.04) | -15.0858* | 1 | 1991q4(0.04) | -8.5488* |
| Unemployment Rate | 0 | 1993q2(0.12) | -6.7359* | 0 | 1993q2(0.12) | -6.7359* |
| *Innovational Outlier Mixed Model* | | | | | | |
| GDP | 1 | 1995q4(0.05) | -4.1245* | 5 | 1995q4(0.05) | - |
| M2 | 0 | 1999q3(0.70) | -7.8329* | 0 | 1999q3(0.70) | -7.8329* |
| M1 | 0 | 1999q3(0.70) | -14.2015* | 4 | 1999q3(0.70) | -6.3960* |
| Deposit Rate | 3 | 2004q1(0.84) | -8.5023* | 3 | 2004q1(0.84) | -8.5023* |
| CPI | 0 | 1973q2(0.08) | -12.6914* | 7 | 1973q2(0.08) | -5.0569* |
| Employment | 4 | 1991q4(0.04) | - | 4 | 1991q4(0.04) | - |
| Unemployment Rate | 9 | 1993q2(0.12) | - | 9 | 1993q2(0.12) | - |

Note: * and ** denote a rejection of the unit root null at 5% and 10% respectively. k refers to the number of the lags of the differences of the dependent variable. In the empty cells the code returned NaN.

**Table 2.11: Qu and Perron (2007) VAR Estimation**

| *VAR* | *SupLR Test* | | | *Break* | | |
|---|---|---|---|---|---|---|
| | *Conditional Mean* | *Variance* | *Conditional Mean and Variance* | *Conditional Mean* | *Variance* | *Conditional Mean and Variance* |
| DGDP, DM1, DDeposit Rate, DCPI | 30.351* | 48.833*** | 87.942*** | 2003q1 | 2003q1 | 2003q1 |
| DGDP, DM2, DDeposit Rate, DCPI | 34.619** | 43.070*** | 71.672*** | 2000q1 | 2000q4 | 2000q1 |
| DGDP, DM1, DDeposit Rate, DCPI, DEmployment | 45.861*** | 66.880*** | 109.448*** | 2000q4 | 2000q4 | 2000q4 |
| DGDP, DM2, DDeposit Rate, DCPI, DEmployment | 59.869*** | 65.441*** | 121.450*** | 2000q1 | 2000q1 | 2000q1 |
| DGDP, DM1, DDeposit Rate, DCPI, DUnemployment | 46.819*** | 54.657*** | 109.424*** | 2003q1 | 2003q1 | 2003q1 |
| DGDP, DM2, DDeposit Rate, DCPI, DUnemployment | 45.352*** | 50.203*** | 97.029*** | 2000q1 | 2000q1 | 2000q1 |

Note: ***,**, and * denote a rejection of the no break null (with H1: one break) at 1%, 5% and 10% respectively.

**Table 2.12: Qu (2007) Co-integration Test**

| *VAR* | *Ho: One Stable Co-integrating Vector* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | *SQ* | | | *WQ* | | | *SupQ1 Test (One Break)* | *SupQ2 Test (Two Breaks)* |
| | 1% | 5% | 10% | 1% | 5% | 10% | | |
| GDP, M1, Deposit Rate, CPI | 7.7132 | 7.5819 | 7.5481 | 4.0020 | 3.8707 | 3.8369 | 3.7111 | 5.8420 |
| GDP, DM2, Deposit Rate, CPI | 6.9773 | 6.8654 | 6.8366 | 3.5665 | 3.5665 | 3.5665 | 3.5665 | 4.9788 |
| GDP, M1, Deposit Rate, CPI, Employment | 10.6562 | 10.4800 | 10.4348 | 5.8642 | 5.8642 | 5.8642 | 5.8642 | 6.9404 |
| GDP, M2, Deposit Rate, CPI, Employment | 10.3002 | 10.1355 | 10.0933 | 5.8210 | 5.8210 | 5.8210 | 5.8210 | 6.4875 |
| GDP, M1, Deposit Rate, CPI, Unemployment | 9.6651 | 9.5003 | 9.4581 | 5.1822 | 5.1822 | 5.1822 | 5.1822 | 6.4929 |
| GDP, M2, Deposit Rate, CPI, Unemployment | 9.5894 | 9.4195 | 9.3760 | 4.9684 | 4.9684 | 4.9684 | 4.9684 | 6.6929 |

Note: ***, **, and * denote a rejection of the null at 1%, 5% and 10% respectively. The critical values for the SupQ test (one break) are: 6.93, .8.05 and 10.70 for 10, 5, and 1% respectively. The critical values for the SupQ test (two breaks) are: 11.06, 12.49 and 15.79 for 10, 5, and 1% respectively. The critical values for the SQ test are: 13.28, 15.21, 19.82 for 10, 5, and 1% respectively. The critical values for the WQ test are: 7.63, 8.89, 11.57 for 10, 5, and 1% respectively.

**Table 2.13: VECM Estimation**

| | *Error Correction* | | | | |
|---|---|---|---|---|---|
| | *D(GDP)* | *D(M2)* | *D(DEPOSIT RATE)* | *D(CPI)* | *D(EMPLOYMENT)* |
| CointEq | -0.103* | -0.338*** | -0.916** | -0.211*** | 0.161* |
| | (0.057) | (0.076) | (0.457) | (0.038) | (0.089) |
| Trend | 0.000 | 0.000** | -0.000 | 0.000 | 0.000 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Constant | 0.009*** | 0.029*** | 0.016 | 0.010*** | -0.001 |
| | (0.002) | (0.003) | (0.022) | (0.001) | (0.004) |

| | | *Co-integrating Equation* | | | | |
|---|---|---|---|---|---|---|
| *GDP* | *M2* | *DEPOSIT RATE* | *CPI* | *EMPLOYMENT* | *Trend* | *Constant* |
| 1.000 | -0.143* | -0.006 | 3.512*** | -0.197* | -0.028 | -31.052 |
| | (0.081) | (0.018) | (0.398) | (0.105) | | |

Note: ***, **, and * denote significance at 1%, 5% and 10% respectively.

## Figure 2.1: First-Difference Presentation of the Variables

*Figure 2.1.1: Gross Domestic Product (GDP)*

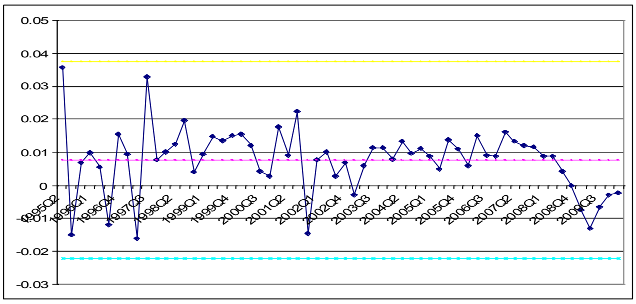


*Figure 2.1.2: Money Supply (M2)*

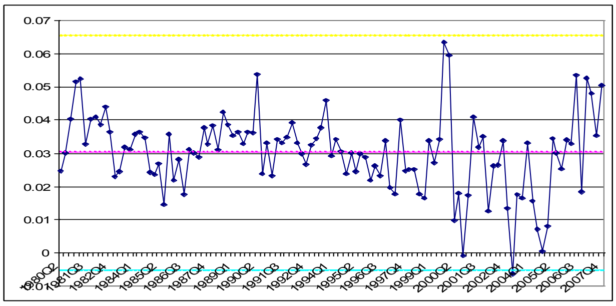


*Figure 2.1.3: Money Supply (M1)*

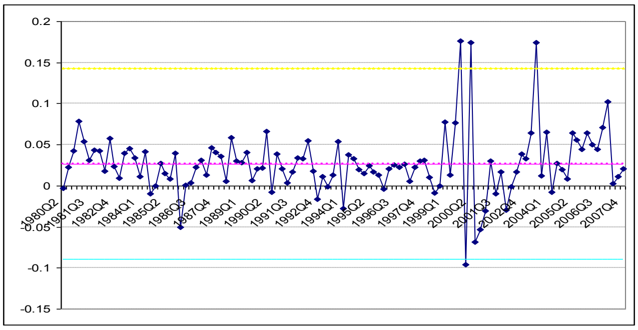


*Figure 2.1.4: Deposit Rate*

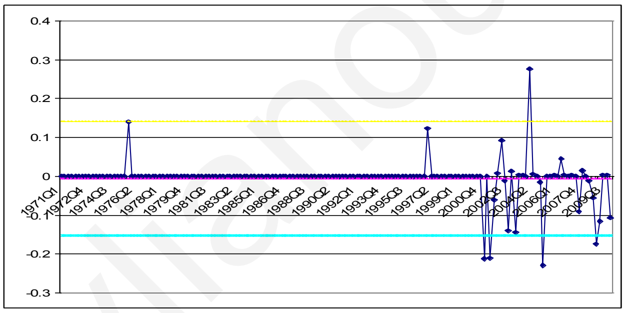


*Figure 2.1.5: Consumer Price Index (CPI)*

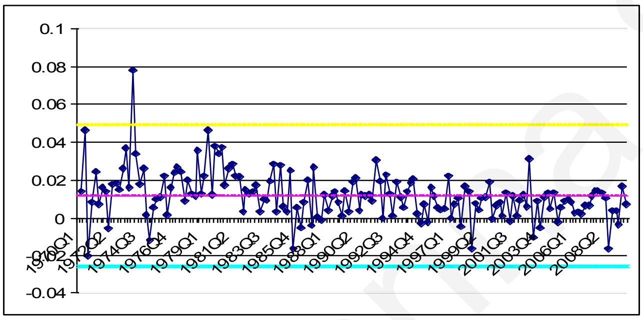


*Figure 2.1.6: Employment*

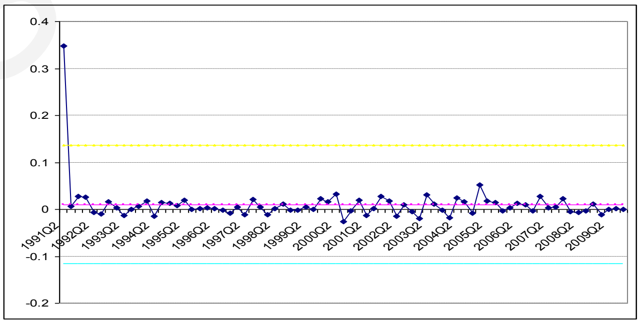


*Figure 2.1.7: Unemployment Rate*

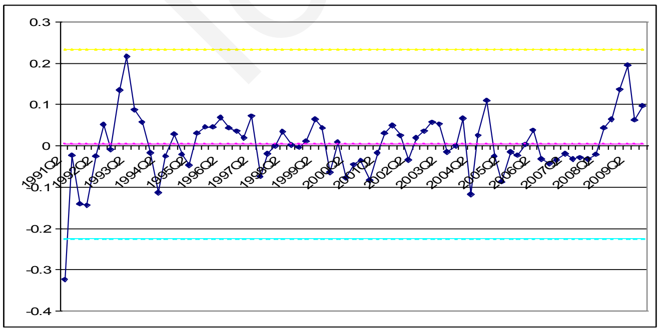
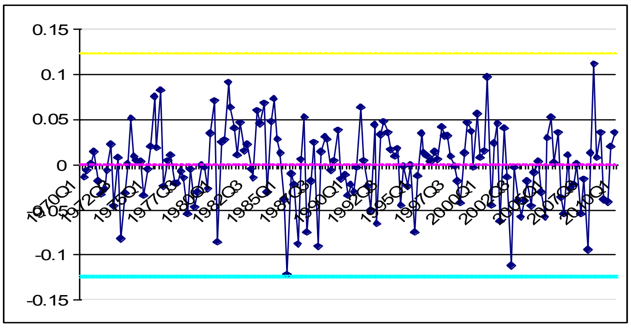
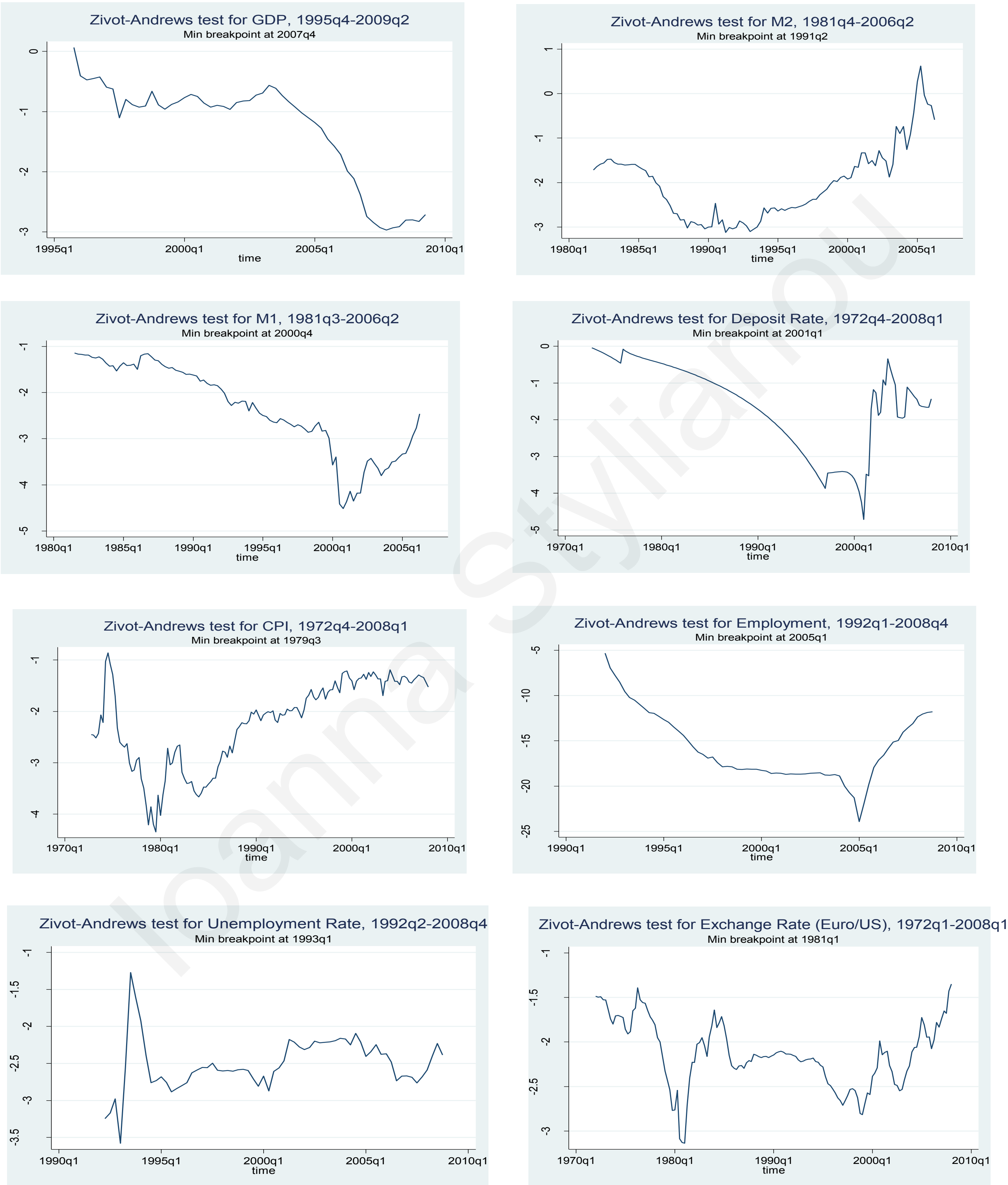


*Figure 2.1.8: Exchange Rate (CYP/USD)*

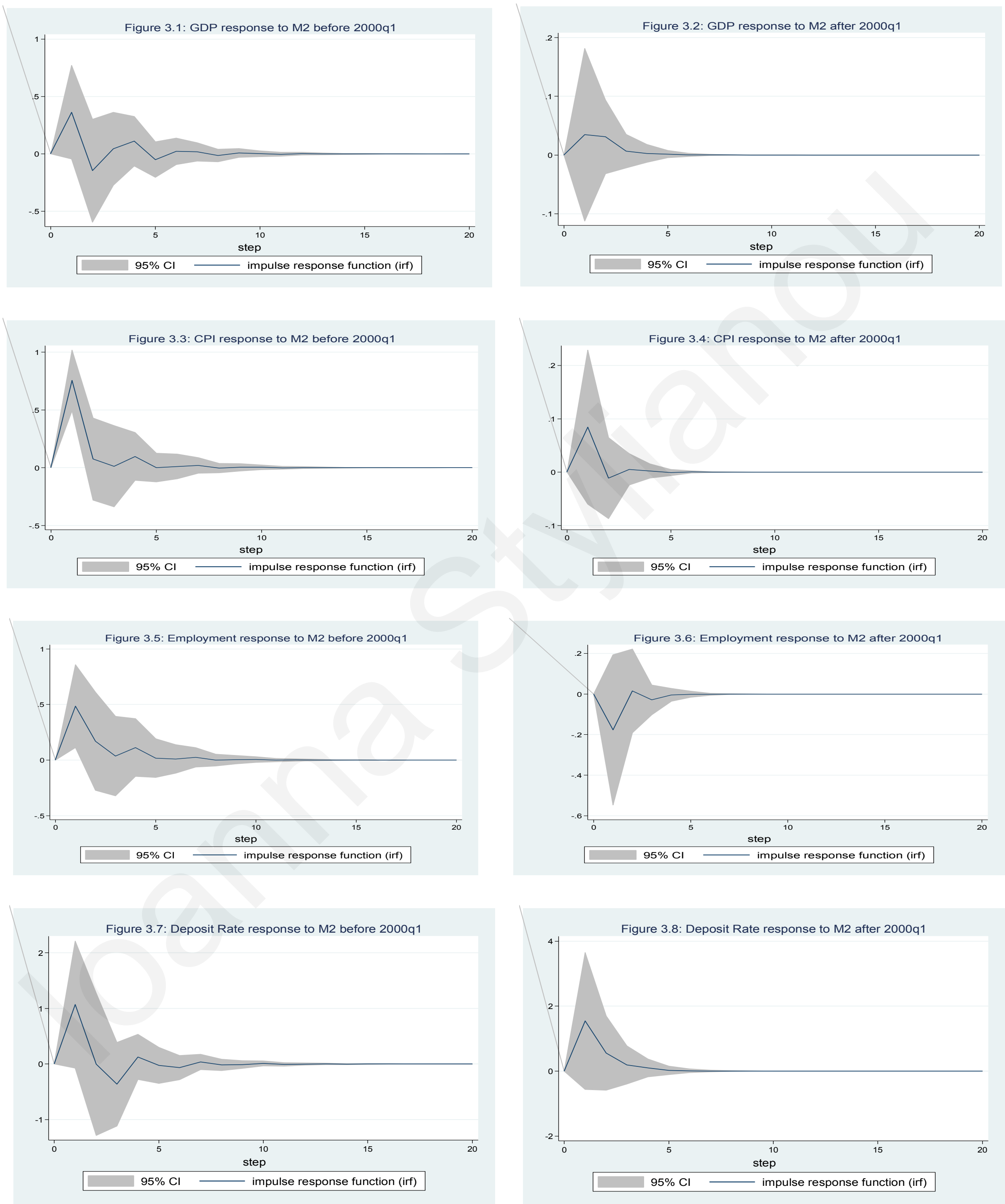


*Note: The value of the upper and the lower bound line is calculated as Mean ± 3xStandard Deviations*

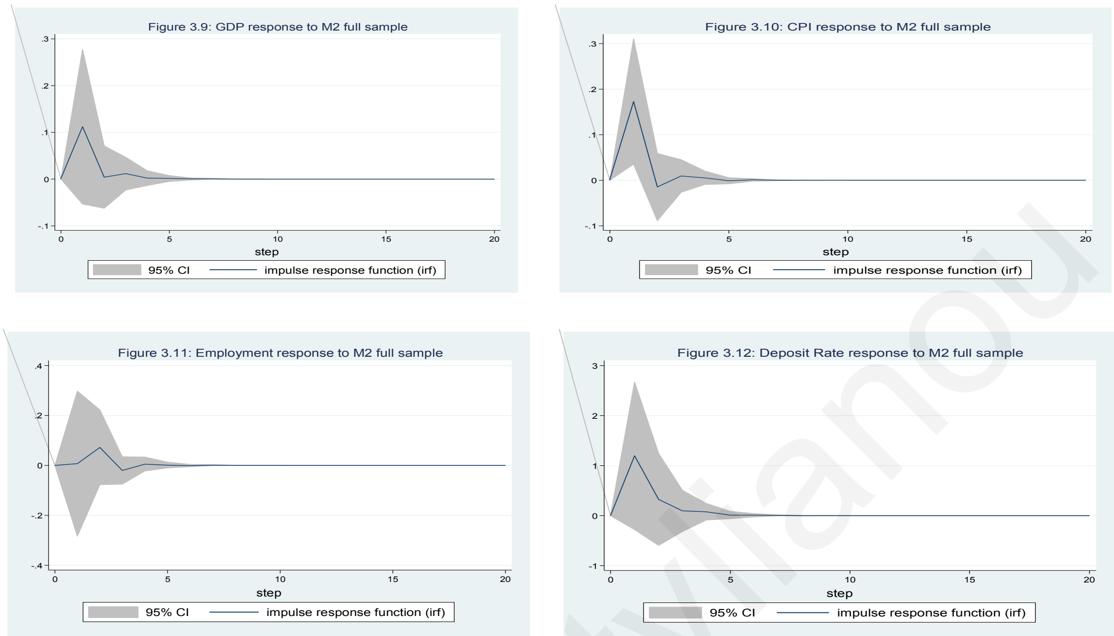
**Figure 2.2: Zivot and Andrews (1992) Min T-Statistic for a Unit Root (Model C)**



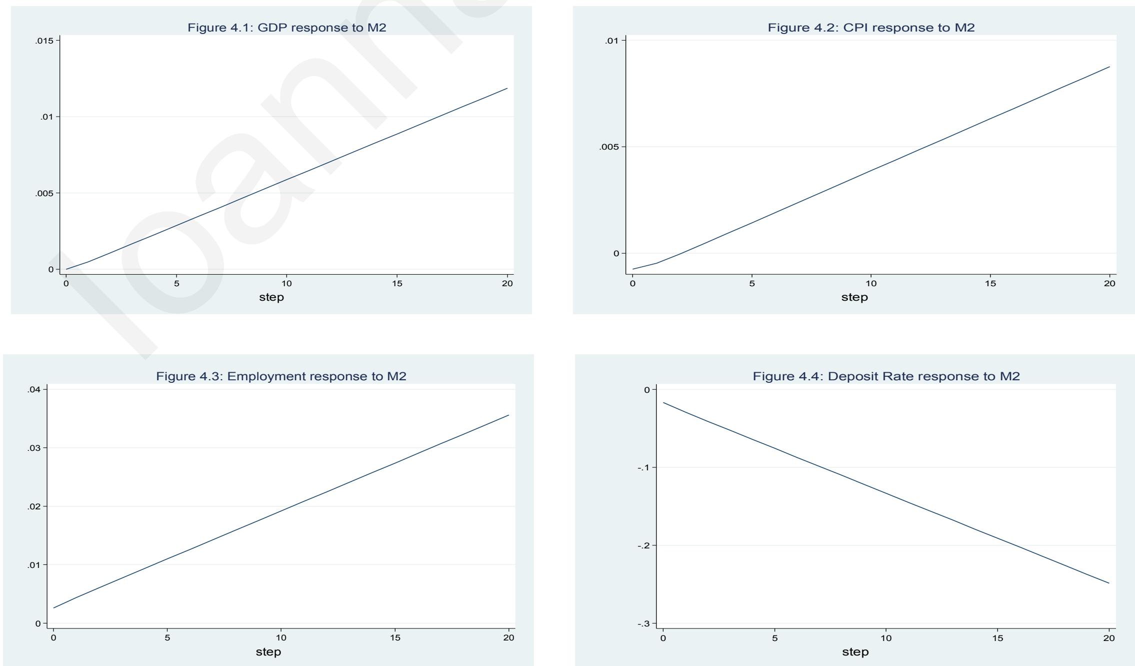
# Figure 2.3: VAR-Impulse Response Functions



## Figure 2.3: VAR-Impulse Response Functions-*Continued*




## Figure 2.4: VECM- Impulse Response Functions



# Chapter 3

## Failure to Launch? The Role of Land Inequality in Transition Delays

## 3.1 Introduction

The transition from economic stagnation to growth and the associated phenomenon of the great divergence has been the subject of intensive research in the growth literature. In particular, there has been a large body of work that is concerned with the issue of economic take-offs. This work describes the transition of economies from a state of economic stagnation to a modern industrial economy with positive growth rates. Notable examples include Galor and Weil (2000), Hansen and Prescott (2001), and Desmet and Parente (2009).

Recent work in the growth literature has focused on the effect of fundamental theories (associated with slow moving determinants) such as geography and institutions on variations in long-run economic performance across countries. We are interested in whether these fundamental determinants are also important explanations for delays in countries achieving economic take-offs. Specifically, this paper focuses on the empirical support for one theory of transition delays - initial land inequality.

Many researchers have highlighted the role that initial land inequality plays in terms of delaying the onset of economic take-off. In particular, the theory has highlighted the deep connection between land inequality and human capital accumulation. In Galor, Moav, and Vollrath (2009), henceforth GMV, land inequality negatively affects the implementation of educational reforms that lead to the extension of educational opportunities to the general population.[1] In particular, due to the low complementarity of human capital and land (see also, Galor and Moav (2006)), an increase in the level of human capital increases productivity in industry more than the agricultural sector, causing a decrease in the returns to land and a rise in wages. Consequently, political elites who initially derive most of their income from land have no incentive to support educational reforms. However, since productivity growth in the industrial sector outstrips that in the agricultural sector, the returns from the capital holdings of political elites increase as a proportion of their total income as the economy advances. Their objection to education reform therefore declines over time such that a critical time is reached whereby human capital-enhancing policies (e.g., compulsory schooling) are enacted.

While GMV posit a direct effect of land inequality on transition delays, other work in the

[1] Several other works have also documented the relationship between land inequality and the lower provision of other forms of public goods (including financial development), such as Banerjee and Iyer (2005) and Rajan and Ramcharan (2010).

literature also propose an indirect effect whereby land inequality influences the evolution of political institutions, and it is these institutions that then determine the delays in transition. Parente and Prescott (2000), Acemoglu, Johnson, and Robinson (2006), and Engerman and Sokoloff (2002) have all pointed out the important role that land inequality plays in determining the evolution of political institutions. The difference between these works and that of GMV is the emphasis on an independent role for political institutions and their persistence in determining delays in enacting human capital promoting initiatives. For example, as Acemoglu, Johnson, and Robinson (2006) point out, if there are rents to staying in power, then, the politically powerful landed aristocracy would have a strong incentive to block the introduction of new technologies and institutions in order to protect their power and profits, delaying at the same time the industrialization process. The suggestion here is that the autonomous nature of political institutions may require direct reforms to these institutions in order for welfare enhancing outcomes to be achieved. In contrast, in GMV's framework, economic progress automatically leads to a shift in incentives faced by the elites, and to their willingness to adopt human capital enhancing policies.

Our contribution in this paper is threefold. First, we ask the question of what factors determine the delay of a country in achieving a particular educational threshold (e.g. 50% primary schooling enrollment). Specifically, do higher levels of land inequality lead to longer delays? We exploit a new historical data set for land inequality by Frankema (2009) to investigate this question in the context of hazard rate models. This is a departure from the standard empirical work that is carried out in the growth literature. Methodologically, empirical work in the growth literature focuses on the effects of various covariates on long-run per capita income or growth. In this paper, we focus instead on a more direct prediction of the theory - what are the effects of various fundamental determinants on delay in schooling?

In addition, we explicitly address the issue of model uncertainty in investigating how fundamental determinants, such as land inequality, affect the extension of schooling opportunities. Our analysis does not assume that the GMV theory is necessarily the true one but rather it provides findings that are robust to alternative theories and their proxies. More precisely, we employ a Bayesian model averaging technique that aggregates the findings across different plausible model specifications using the posterior evidence as weights for each model; see for example Durlauf, Kourtellos, and Tan (2008).

Consistent with the theory proposed by GMV, we find that increased levels of land inequality lead to more delays in reaching the 50% primary schooling enrollment rate threshold. This result is robust to variations in the specification of the hazard model, and also holds true for other primary schooling threshold levels that are consistent with a substantial extension of schooling opportunities to the population. Interestingly, initial values of political institutions (as measured by an executive constraint variable) do not appear to be important in determining delays in schooling.

Second, we explore the question of how delays in schooling affect contemporaneous measures of schooling and institutions. We find strong evidence that historical delays in achieving particular educational thresholds determine not only current schooling but also current executive constraints. This evidence suggests an alternative channel for the effect of land inequality on long-run economic performance via schooling delays. Therefore, our third contribution investigates whether schooling delays are transmitted to long-run income via their effect on current institutions and schooling. Like Glaeser, La Porta, Lopez-De-Silanes, and Shleifer (2004), our results suggest that the historical evolutionary path of human capital is a fundamental determinant of long-run economic outcomes.

The paper is organized as follows. Section 3.2 lays out the econometric framework and discusses our data and results for the hazard analysis. Next we investigate the potential implications of our findings for current schooling, institutions, and income in Section 3.3. Section 3.4 concludes.

## 3.2 Econometric framework

### 3.2.1 Implementation and data

We employ the static Cox (1972) proportional hazard (Cox-PH) model to study the probability of the event that a country moves from a low education state to a high education state. In GMV's theory, these two states correspond to a state of economic stagnation and a state of sustained economic growth, respectively. Higher hazard rates correspond to higher risks of transitioning out of the low education/stagnation state therefore implying shorter delays in the extension of schooling opportunities.

The Cox-PH model decomposes the hazard function into a part that depends on the time already spent in stagnation and on a set of explanatory variables $X_i$. So we can write the hazard function for country $i$ at time $t$ as

$$\lambda(t|X_i;\theta) = \frac{f(t|X_i;\theta)}{S(t|X_i;\theta)} = \lambda_0(t)\exp(X_i'\theta), \tag{3.2.1}$$

where $\lambda_0(t)$ is the time dependent part also known as the baseline hazard function. We estimate (3.2.1) using a partial likelihood method. The Cox-PH model is a semi-parametric model in the sense that while it makes no assumption about the form of the function $\lambda_0(t)$, it assumes parametric form for the effect of the predictors on the hazard.[2]

In terms of the data, we construct a historical dataset spanning from 1700 to 1995 for a sample of 53 countries. A detailed description of the data and our sources is given in the Data Appendix B1.

The dependent variable in this case is the delay in schooling, measured as the time it takes for each country to first reach a threshold level in primary schooling enrollment, minus the time it took the first country to pass that threshold. For example, the United States was the first country to pass the 50% primary school enrollment threshold in 1831. The specific variable is close to the GMVs theory and captures the idea of effectiveness of public education. [3]

The reason for constructing the delay variable as a measure that relates primary schooling enrollment in one country relative to the first country to pass the threshold is so as to overcome the left censoring problem. Since all other countries achieve the threshold at later dates than the first country to do so, left censoring is eliminated. The threshold levels we consider have to fulfill two conditions: (i) they have to be high enough to capture the GMV idea of a large scale extension of public schooling opportunities to the population, and (ii) they have to be low enough so that enough countries attain the level within the sample so that we do not have too many instances of right censoring in the data.

The actual construction of the primary schooling data follows Comin and Hobijn (2004) who construct historical primary schooling data for 23 industrialized countries, measured as the number of students in primary school as a fraction of the population, in the age range 5-



[2]For robustness purposes we also investigated parametric methods such as the Exponential, Log Logistic, and Weibull without finding substantial differences.

[3]Alternatively, we could use a dummy 0 or 1 when a country introduces public education or public expenditure as a continuous variable or years of schooling. However, we choose schooling delays because it is closer to the related theory.

14. We extend the primary schooling data set to a larger set of 53 countries for this analysis. The number of students enrolled is taken from Banks (1999), while the population in the age range 5-14 is taken from Mitchell (1998). Table 3.1 lists the countries in our sample.

We now discuss the set of explanatory variables, $X_i$. For our analysis to correspond closely with the theory, we imagine that countries always existed, but have different structural characteristics and historical experiences that influence when they achieve a particular threshold level in schooling. These factors then explain why a particular country experienced a delay in schooling attainment vis-à-vis the US experience. We think of these factors as controlling for two kinds of country-specific heterogeneity.

The first type of country-specific heterogeneity corresponds to factors that are invariant to the particular political elites that are in power at the time when schooling policy decisions are made. These factors largely correspond to country-specific fixed effects as well as the time it took for the relevant political elites; that is, the political elites who would make policy decisions about schooling and who would see these through, to come to power.

One reason why a country might have experienced a delay in schooling attainment vis-à-vis the US may be because of its colonial history. We do have information about whether a country was historically a European colony. To the extent that we can think of the initial conditions of a colony as being substantially influenced by the European metropolis, we can control for country-specific heterogeneity by including colonial dummy variables (specifically, whether a country was a British colony, a Spanish or Portuguese colony, a French colony, or Other European colony).

Another reason why a country may take more time than the US to attain a particular schooling threshold may be that the relevant elites took longer to attain power and therefore control over schooling policies. To control for the variation in the time it took a country's elites to attain autonomy over policies relative to the US, we include an Independence variable that measures the additional years it took for each country to declare independence relative to the US, who declared independence in 1776. This variable takes the value zero for metropolis countries, and positive integers for colonies.

We also control for the elites' hold on power by including a measure of Political Instability due to Miller (2011). The idea is that elites who cannot secure their hold on power may have less ability to influence policy outcomes (or, alternatively, face different incentives in enacting particular policies) hence leading to variation in delays in achieving particular schooling thresholds. Political Instability is measured as the average of the first differences (in absolute values) of the Polity2 variable from Polity IV. The Polity2 variable is a measure of the degree of democracy in a country with a score of +10 representing most democratic and -10 signifying most autocratic. The averages of the first differences are calculated as follows: for colonies we average values of the (absolute) year-to-year changes in the Polity2 variable from the year of independence to the year the colony achieves the schooling threshold, while for non-colonies, we take the corresponding average values from the earliest available observation until the year the country achieves the schooling threshold.

The second set of variables corresponds to factors that influence the incentives of political elites to extend primary schooling opportunities to the population according to the theory. As detailed in the Introduction, our main aim is to investigate how land inequality affects the transition from economic stagnation to the sustained growth era, through the human capital channel. To do so, we use land inequality data from Frankema (2009). The variable is expressed in Gini coefficients, and it is compiled on the basis of the decile distribution of the total number of land holdings (farms), and the total amount of agricultural land (nation-wide), excluding communal pastures and forests.[4] Here, a holding refers to "all agricultural land assigned to a "holder" that is one or two persons, but no group, community or state, or to a distinct "management unit", i.e. a farm. The total agricultural area includes all land that is part of a holding, i.e. arable land, land under permanent crops, land under permanent meadows and pastures, wood and forest land and a category of all other land. In the case of shifting cultivation the total area of the holding consists of the total area under crops and the area that is prepared for cultivation [Frankema, 2006, p. 3]". The primary data sources that Frankema uses to calculate the land distribution data comes from the IIA and FAO World Census of Agriculture. For our analysis, we use the earliest available land Gini observation for each country.

One concern with using land Gini as a variable is that it may be proxying for other forms of wealth or income inequality. Some forms of wealth inequality may in fact imply dramatically different theoretical outcomes from those of GMV. For example, if inequality was a result of inequality in capital holdings and not of land holdings by elites, it may be the case that elites would prefer higher levels of schooling for the population since human capital is complementary to physical capital. However, if we fail to include a proxy



[4] The land Gini coefficient is defined as, $G = \sum_{i=1}^{n} \sum_{j=1}^{n} |z_i - z_j| \ / \ 2n^2\mu$, where $z_i$ and $z_j$ are the percentage shares of land of $n$ deciles ($n = 10$) and $\mu = 1/n$.

variable for capital holdings inequality then the estimates for the effects of land inequality on schooling outcomes are likely to be biased. Alternatively, the precise nature of the inequality responsible for lower schooling levels may be misspecified. For example, land Gini may be proxying for income inequality (instead of land inequality) which has also been shown to be associated with poor education outcomes across countries. In fact, Goldin and Katz (1997) find evidence that supports this proposition for the case of the US. To safeguard against these possibilities, we obtain data for the number of automobile registrations per person to proxy for other forms of inequality not associated with land. We use the earliest available value for automobile registration for each country and note that all values were taken at dates that preceded the year in which the country reached the schooling threshold that defines the dependent variable. However, the use of per capita automobile registration is potentially a very imperfect measure of income inequality. Historically, it is probably most informative only in the very particular historical situation of the early 1900's in the US or Europe, and may therefore not be a very good measure of inequality for other regions of the world. We therefore consider a new dataset of global inequality (BFLZ Gini Index) that has recently been introduced by Van Zanden, Baten, Földvari, and Van Leeuwen (2011). This new dataset is available for a large set of countries spanning from 1820 to 1995 and improves the Bourguignon and Morrisson (2002) dataset in several ways. In particular, it is calculated using a much larger number of observations of within country inequality and it is based on the new 2005 PPPs of the World Bank's ICP project, which gives a more accurate picture of disparities in GDP per capita than the previous ICP rounds.[5]

GMV also theorize that land abundance that would benefit agriculture in the early stages of development would lead the landowning elite to be more reluctant to enact human capital enhancing policies that disproportionately benefit capitalists and workers. We include therefore a measure of land abundance, the log of arable land (absolute) in hectares, in 1700, and investigate whether more land abundance leads to greater delays in schooling.

GMV's theory also requires that we control for other developmental differences between countries. The reason is that for a given level of land inequality, all else equal, higher levels of economic development corresponds to capital holdings constituting a larger proportion of the asset portfolio of elites. Since elites in more developed economies would derive a higher portion of their income from the industrial sector, they would be more willing to enact human

[5]We are very grateful to Bas van Leeuwen for very kindly sharing the global inequality data from Van Zanden, Baten, Földvari, and Van Leeuwen (2011) with us.

capital-friendly policies. We control for initial development differences between countries using the log of GDP per capita (Initial Income; Maddison (2009)). For non-colonies, we take the average of log GDP per capita values from the earliest possible data point until 1831, while for colonies, we use the data on independence day or, if this is unavailable, the earliest data point after independence. We should note that in all cases, the income data occurs prior to the country achieving the schooling threshold. The timing of the variable is meant to capture the level of development that was relevant to the elites that are in power at the time when schooling policies are enacted.

The main alternative theory for schooling delays, as noted in the Introduction, is political institutions. We proxy initial political institutions using historical executive constraints data from Polity IV (Initial Executive Constraints). This variable lies between zero and one, with higher values indicating more constraints on the power of the executive. Similar to the Initial Income variable above, we take the average of executive constraints values from the earliest possible data point until 1831 for non-colonies and use the data on independence day or, if unavailable, the earliest data point after independence for colonies. In all cases, the data for executive constraints occurs prior to the country achieving the schooling threshold, and is meant to capture the relevant degree of executive constraints that apply to elites empowered to determine schooling outcomes.

Following the empirical growth literature, we also control for a set of new growth theories that have potential implications for human capital accumulation. The first such theory argues that a detrimental climate may have negative effects on human capital accumulation (see, Sachs, Gallup, and Mellinger (1999)). We proxy climate using a variable (Tropics) which measures the percentage of a country's land area that is classified as tropical or subtropical. Finally, another theory requires that we account for the effects of ethnic heterogeneity on delays in schooling. Alesina, Baqir, and Easterly (1999) suggest that higher levels of ethnic heterogeneity potentially result in political disagreements over the provision of public goods (such as schooling), and its subsequent under-provision. To control for the effect of ethnic heterogeneity on delays in schooling, we include a measure of ethnic fractionalization due to Alesina, Devleeschauwer, Easterly, Kurlat, and Wacziarg (2003) in $X_i$.

Table 4.1 presents summary statistics while Table B1 of the Appendix provides a detailed descriptions of all the variables.

### 3.2.2 Model averaging of hazard models

Standard duration analysis estimates a baseline PH-Cox model in equation (3.2.1), which is closest to the theory in question and then reports a few robustness exercises that include some additional controls. An alternative approach to evaluate the relative evidentiary support of competing theories includes a large number of variables and those variables that prove to be significant are then rendered as the important determinants. This approach is often referred as a 'kitchen sink' approach.

However, both approaches do not systematically address the problem of model uncertainty and do not provide robust evidence but rather rely on strong priors of the econometrician. As Brock and Durlauf (2001) and others have argued, the inherent open-endedness of new growth theories presents unique challenges to researchers in exploring their quantitative consequences on growth. Because the inclusion of one set of growth theories says nothing about whether other possible growth theories should be included (or not) in the model, growth researchers face substantial model uncertainty in their work. The fear is that the inclusion or exclusion of growth variables may significantly alter the conclusions one had previously arrived at for, say, the effect of land inequality on delays in schooling based on a particular model in the model space. In this case, the model space refers to the set of all possible models generated by the set of growth regressors, denoted by $M = M_1, \ldots, M_K$. How can we obtain robust conclusions about the effect of land inequality in equation (3.2.1) and more generally about the structural parameters $\theta$ that do not condition on the model choice?

We do so by employing a Bayesian Model Averaging (BMA) approach by constructing estimates conditional not on a single model, but on a model space whose elements span an appropriate range of determinants suggested by a large body of work. A number of recent papers have documented the advantages of using BMA in constructing robust estimates primarily in the context of the linear model. See for example, Brock and Durlauf (2001), Fernandez, Ley, and Steel (2001), Sala-I-Martin, X. and Doppelhofer, G. and Miller, R. (2004), Durlauf, Kourtellos, and Tan (2008), and Masanjala and Papageorgiou (2008). Our BMA approach is closest to Volinsky, Madigan, Raftery, and Kronmal (1997) who employ BMA in the context of Cox-PH models to study the risk factors for stroke. Model averaging "integrates out" the uncertainty over models by taking the weighted average of model-specific estimates, where the weights reflect the evidentiary support for each model given the data,

$D$, and which are constructed to be analogous to posterior model probabilities. Then the posterior distribution of $\theta$ given the data, $D$, is given by

$$\widehat{\mu}(\theta|D) = \sum_{k=1}^{K} \widehat{\mu}(\theta|M_k, D)\widehat{\mu}(M_k|D)) \tag{3.2.2}$$

where $\mu(\theta|M_k, D)$ is the posterior distribution of $\theta$ given a particular model $M_k$, and $\mu(M_k|D)$ is the posterior probability of model $M_k$. The former is a standard Bayesian object, which does not have a closed form expressions in the case of Cox-PH models. Following Volinsky, Madigan, Raftery, and Kronmal (1997) we approximate it by the maximum likelihood estimator, $\widehat{\mu}(\theta|M_k, D) \approx \widehat{\mu}(\theta|M_k, \widehat{\theta}_k^{MLE}, D)$.[6]

As for the model weights, $\widehat{\mu}(M_k|D)$ we use the Bayes' rule, so that each weight is the product of the integrated likelihood of the data given a model, $\widehat{\mu}(D|M_k)$, and the prior probability for a model, $\mu(M_k)$:

$$\widehat{\mu}(M_k|D) \propto \widehat{\mu}(\theta|M_k)\mu(M_k) \tag{3.2.3}$$

As standard in the literature, we assume a uniform prior so that the prior probability that any variable is included in the true model is taken to be 0.5. The integrated likelihood of model $M_k$ is approximated by the Bayesian information criterion (BIC), $\log \widehat{\mu}(D|M_k) = \log \widehat{\mu}(D|\widehat{\theta}_k^{PLE}, M_k) - (p/2)\log n + O(1)$, where $n$ should be the total number of uncensored cases.

The model averaging estimator of $\theta$ is given by the posterior mean defined by

$$\widehat{\theta}_{D,M}^{MA} = \sum_{k\in M} \widehat{\theta}_k^{PLE}\widehat{\mu}(M_k|D), \tag{3.2.4}$$

where $\widehat{\theta}_k^{PLE}$ is the partial likelihood estimator of each model $M_k$. We compute the corresponding standard errors using the posterior variance of $\theta$

$$\widehat{V}_{D,M}^{\theta} = \sum_{k\in M} \widehat{V}_{D,k}^{\theta}\widehat{\mu}(M_k|D) + \sum_{k\in M}(\widehat{\theta}_{D,k}^{PLE} - \widehat{\theta}_{D,M^S}^{MA})^2\widehat{\mu}(M_k|D) \tag{3.2.5}$$

[6]This posterior refers to the following integral $\mu(\theta|M_k, D) = \int \mu(\theta|\theta_k, M_k, D)\mu(\theta_k|M_k, D)d\theta_k$, where $\mu(\theta|\theta_k, M_k, D)$ is the likelihood and $\mu(\theta_k|M_k, D)d\theta_k$ is the prior density of $\theta_k$.

where $\widehat{V}_{D,k}^{\theta}$ is the model-specific posterior variance of the partial likelihood estimator estimator. The first term in equation (4.2.9) is the average of the posterior variances within models and the second term is the variance of the posterior means across models (i.e. weighted average of the squared deviations of the model-specific from the model averaged estimates).[7] We also report the posterior probability of inclusion for each covariate, which is the sum of the posterior probability of all the models for which that variable appears. It is meant to capture the (posterior) probability that that covariate is in the true model after looking at the data. A way of accounting for model uncertainty involves applying the Occam's Window algorithm of Madigan and Raftery (1994) to linear regression models. Two basic principles underly this approach. First, if a model predicts the data far less well than the model which provides the best predictions, then it has effectively been discredited and should no longer be considered. Second, appealing to Occam's razor, models which receive less support from the data than any of their simpler sub-models, are excluded. This greatly reduced the number of models in equation 3.2.2.

### 3.2.3 Hazard results for delay in schooling

We present our findings for the Cox PH model in equation (3.2.1) in Table 3.4. The dependent variable, delay in schooling, is the time it takes for each country to first reach 50% in primary schooling enrollment, minus the time the first country (the US, in this case) passed the 50% threshold. We chose to focus on the 50% threshold for two reasons.

First, the 50% threshold level is an appropriate level as it is consistent with the GMV idea of a substantial extension of schooling opportunities to the population. However, to get a sense of the robustness of our findings, we also investigated various other schooling threshold values ranging from 40% to 60% primary schooling threshold levels. The 50% threshold level also turns out to be neither too high nor too low in the following sense. When the threshold level is low (essentially for all threshold levels below 45%), almost all countries successfully attain the threshold level with very little difference in the time it took to do so, so that there is not enough variation in the data to properly identify the effects of land inequality on schooling delays. However, when the threshold level is high (above 55%), the number of right censored countries becomes large. Table 3.3 shows the countries

[7] Our approach can be viewed as a "hybrid" approach to model averaging in the sense that we mix frequentist probability statements about observables given unobservables and Bayesian probability statement about unobservables given observables. For a similar approach, see Durlauf, Kourtellos, and Tan (2011)

that failed to reach various primary schooling enrollment threshold levels; i.e., countries that are right censored. Right censoring reduces the observed variation in schooling delays, and makes it difficult to identify the effects of land inequality on delays.[8]

The first three columns of Table 3.4 present the results from our model averaging analysis. The first column shows the posterior probability that each of the covariates is included in the true model for the hazard rate, while the second and third columns present the BMA posterior means and standard errors for each covariate. The remaining six columns show, respectively, the coefficient estimate and standard error for each covariate for (i) the two posterior mode models from the BMA analysis, and (ii) the largest model in the model space considered in the BMA analysis.

Our reason for reporting the results from the posterior mode and largest models is to provide the reader with the ability to compare findings via model selection - using the best models (in terms of posterior weights) or a low-bias model (at the cost of reduced efficiency) with potentially many irrelevant covariates - with those obtained via model averaging (BMA). Finally, we also note that the posterior means are interpreted as the marginal effect of each covariate on the risk of crossing the 50% primary schooling threshold. Therefore, positive estimates imply that the marginal contribution of the corresponding covariate is to reduce the delay in schooling for countries.

Our BMA results are consistent with the theoretical predictions of GMV. As GMV argued, for given levels of economic progress, land inequality implies a higher reliance of political elites on income derived from landholdings leading them to delay the implementation of human capital enhancing policies, which primarily benefit capitalists and workers. Similarly, the greater the abundance of arable land, all else equal, the greater the importance of agriculture in the elites' portfolio, the higher their subsequent reliance on returns from landholdings, and the greater their reluctance to expand schooling opportunities. However, for given levels of land inequality and arable land, economic progress results in a rebalancing of the portfolio returns of landholding elites away from income derived from land holdings to returns from capital holdings resulting in elites being more willing to extend schooling to the population.

[8] For conciseness, we only report full results for the 50% threshold. The results for land gini as well as the other covariates for threshold levels between about 40% and 50% do not differ substantively. This can be seen from Appendix Figure A1, which shows the Posterior Inclusion Probabilities for the land gini variables across threshold levels. Full results for all other covariates are available upon request.

Consistent with the theory, we find that higher levels of land inequality (higher values for Land Gini), greater abundance of Arable Land, and lower Initial Income result in lower risks of exceeding the 50% schooling threshold, thereby implying greater delays in the expansion of schooling opportunities. More precisely, the posterior inclusion probabilities of Land Gini, Arable Land, and Initial Income are all very high at 98.9%, 97.8%, and 100%, respectively - well above the 50% prior inclusion probability. The corresponding posterior means for all three variables are also strongly significant at the 1% level. The BMA results are confirmed by the results from the posterior mode models. Accounting for model uncertainty by averaging across models delivers the same conclusions as doing so by selecting (the best) models.

Table 3.4 also makes clear that it is inequality in land ownership specifically, and not other (non-land) forms of inequality (as proxied for by Auto Registration and BFLZ Gini Index) that is important in determining schooling delays. The posterior inclusion probability for both Auto Registration and BFLZ Gini Index are negligible at 7.2% and 11.9% respectively. The posterior means for both variables are also not significant. Finally, neither Auto Registration nor BFLZ Gini Index appears as a covariate in either of the two posterior mode models.[9]

Two sets of factors that can be interpreted as country fixed effects are shown to be strongly significant. The first is the delay in a country gaining independence relative to the US. We find, predictably, that countries that took more time to gain independence, so that the relevant elites required more time to attain autonomous control over policies, also faced longer delays in achieving an extension of schooling opportunities to the population, all else equal. The posterior inclusion probability of the Independence variable is 100% and significant at the 1% level. Along with gaining autonomy over a country's policies, the level of Political Instability (elites' hold over power) is also important (with posterior inclusion probability of 100%) and highly significant at the 1% level. Our BMA findings (consistent with those of the other reported models) indicate that a greater degree of Political Instability, all else equal, leads to longer delays in reaching the 50% schooling threshold. Finally, being either a British colony or some other colony that is not French, Spanish, or Portuguese results in a shorter delay in achieving schooling take-off. The posterior inclusion probabilities for the British colony and Other colony dummies are both very high at 95.5% and 100%, respectively, and the corresponding coefficient estimates are strongly significant.

[9]In Table A2 of the Appendix, we consider the robustness of our results to dropping either one or both of the inequality variables, Auto Registration and BFLZ Gini Index. We find that our results are generally robust to these alternative specifications.

None of the other growth theories appear to be either significant or important (in terms of posterior inclusion probabilities) explanations for delays in achieving the schooling threshold. Importantly, the results in Table 3.4 make clear that there is no evidence that initial institutions (as measured by Initial Executive Constraints) affects schooling delays. The posterior inclusion probability for Initial Executive Constraints is well below 50% at 15.2%, and the posterior mean is not significant. Initial Executive Constraints also does not appear in either of the posterior mode models.

Finally, the two posterior mode models are very similar (they differ only in the inclusion/exclusion of Ethnic Fractionalization) and have posterior model probabilities that are very close (0.194 and 0.175). Outside of these two models, the other models in the model space have very low posterior probability, and are therefore not reported. For example, the third best model in the model space has negligible posterior model probability (0.076), and the largest model has posterior model probability of 0.000.

In sum, our findings appear to provide strong support for the hypothesis that schooling delays are entirely explained by variables suggested by GMVs theory.

## 3.3 Long-run implications of delays in take-offs

### 3.3.1 Current institutions and schooling

We next turn to the question of whether land inequality has long-term implications for economic performance via its influence on schooling delays. We do this in two steps. In this section, we take the first step by examining whether (historical) delays in the extension of public schooling generate persistent outcomes in current schooling and current institutional quality of a country. In the next section, we will take the next step and investigate the influence of (historical) schooling delays on long-run per capita income via its effect on the (current) measures of schooling and institutional quality considered in this section.

There is general agreement in the literature that the processes of institutions and schooling are highly persistent (see for example Acemoglu, Johnson, and Robinson (2001) and Glaeser, La Porta, Lopez-De-Silanes, and Shleifer (2004)). However, there is disagreement over the the role of other historical factors in determining current institutions and schooling. In a seminal paper, Acemoglu, Johnson, and Robinson (2001) argued that the mortality rate of European settlers in the colonies was the key factor that determined their decision to settle. Since these early European colonizers were more likely to establish higher quality institutions in lands in which they chose to settle, they thereby influenced the formation of early institutions in the colonies. The effect of these early institutions was thought to be persistent, so that these initial/historical institutions became important determinants of the current institutions of a country.

However, Glaeser, La Porta, Lopez-De-Silanes, and Shleifer (2004), using a limited schooling sample from 1960-2000, produced evidence that suggests an alternative channel through which early European settlers influenced the development of subsequent institutions. Glaeser et al suggest that what European colonizers brought with them to the colonies they settled were not the institutions from their home countries, but rather the high levels of human capital that they possessed. It is this early human capital that was responsible for sustaining the quality of institutions of a country and allowed the latter to persist over time.

More recently, Gallego (2010) has argued for the reverse. Gallego considers human capital accumulation to be a consequence of the development of democratic political structures. He appends the story by Acemoglu, Johnson, and Robinson (2001) by hypothesizing that European colonizers who chose to settle in a location were more likely to invest in human capital for their children and for the native population, while those who set up extractive states would have very little reason to do so. Gallego finds that institutions are responsible for current (as well as historical) schooling levels.

Our analysis of the influence of land inequality on long-run outcomes via its effects on schooling delays allows us to revisit the debate over the historical determinants of current institutions and schooling. With the exception that our primarily interest is in documenting the effects of land inequality on contemporaneous outcomes, the perspective we adopt is not very different from the papers cited above. Different levels of land inequality resulted in variations in the delay in which countries achieved large scale extension of schooling opportunities. If the process for schooling is persistent, then, we should find that current levels of schooling are influenced by historical delays in achieving particular schooling thresholds. If human capital accumulation is required for sustaining high quality of institutions, then, we should also observe that shorter delays in achieving particular thresholds of human capital levels in the past should correlate with better quality institutions now.

To address this question, we consider the regressions of current institutions, $R_{T,i}$, and current schooling, $H_{T,i}$, in equations (3.3.1) and (3.3.2), respectively.

$$R_{T,i} = \mu_R + \alpha_R \lambda_i + Z_i' \beta_R + e_{R,i}, \tag{3.3.1}$$

$$H_{T,i} = \mu_H + \alpha_H \lambda_i + Z_i' \beta_H + e_{H,i}, \tag{3.3.2}$$

We measure current institutions, $R_{T,i}$, using average executive constraints over the periods 1965- 1995 and 1985-1995. Our preferred measure for averages executive constraints is the period 1985-95, which is the same period average that Acemoglu, Johnson, and Robinson (2001) use for their institutions measure, and therefore allows our work to be more comparable with the findings in the existing literature. For robustness, we also include results where executive constraints are averaged across the period 1965-95; that is, the period of time after decolonization. Similarly, we measure current schooling, $H_{T,i}$, using the logarithm of average years of male secondary and higher school attainment over the periods of 1965-95 and 1985-95.

Our key determinant of both current schooling and executive constraints is the log hazard rate (Log Hazard), $\lambda_i$. The Log Hazard captures the effect of schooling delays on outcome variables and is defined as $\lambda_i = log(\lambda(t|X_i;\theta)/\lambda_0(t)) = X_i'\theta$, where $\theta$ is estimated by the Cox regression in (3.2.1). $Z_i$ is a vector of additional exogenous control variables, which includes initial values of Schooling and Executive Constraints, Colonial dummies, Tropics, and Ethnic Fractionalization. We also includes proxies for a country's legal system based on British common law (Britcommon), or French civil law (Frecivil) due to La Porta, Lopez-de Silanes, Shleifer, and Vishny (1999). Legal origin, and in particular French civil law, has been found to be an important determinant of both schooling and institutions; see Glaeser, La Porta, Lopez-De-Silanes, and Shleifer (2004) and Acemoglu and Johnson (2005). The coefficients $\alpha_R$ and $\alpha_H$ capture the effect of delay in schooling on current institutions and current schooling, respectively, while $e_{R,i}$ and $e_{H,i}$ are regression error terms.

Tables 3.5 and 3.6 present the results for current Schooling and current Institutions, respectively, over the periods 1965-95 and 1985-95. We focus on the 1985-95 results and contrast them with the 1965-95 results when there is substantive disagreements between the two findings. We present model averaging (2SLS-BMA) results, the posterior mode model, as well as the largest model in the model space. We should note that the BMA methodology here differs from that employed in Section 3.2.3 in that the model averaging estimates refer to weighted sums of 2SLS estimates rather than PLE estimates.[10]

There are two main findings for current Schooling. First, there is no evidence that initial institutions is an important determinant of current Schooling once we control for Log Hazard. The posterior inclusion probability of Initial Executive Constraints from the BMA analysis is far below the 50% prior at 16.6% for the 1985-95 period, and the posterior mean is not significant. The posterior mode model in this exercise has posterior model probability of 11.8% which is slightly more than twice as large as that for the next best model (at 5.2%). Nevertheless, a posterior model probability of 11.8% is not large, and therefore we prefer the BMA results. In any case, Initial Executive Constraints does not appear in the posterior mode model. Second, the only determinant that is both an important (in terms of posterior inclusion probability) as well as a significant determinant of current Schooling is the Log Hazard. The posterior inclusion probability for Log Hazard is 100% and the posterior mean is significant at the 5% level. Hence, land inequality appears to exert a strong influence on current Schooling via its effect on the (historical) delay in schooling.

The main finding for current Executive Constraints is that both initial institutions and schooling delays appear to be important determinants. The posterior inclusion probabilities for both Initial Executive Constraints and Log Hazard are high at 98.7% and 89.7%, respectively. However, for the 1985-95 period, the posterior mean for Initial Executive Constraints is significant at the 5% level while that for Log Hazard is not significant. This result for the posterior mean for Log Hazard, however, appears to be confined to the 1985-95 BMA exercise. The corresponding posterior mean for the 1965-95 period is significant at the 5% level. Also, Log Hazard is a variable that is included in the posterior mode models in both the 1985-95 and 1965-95 exercises. In both these cases, the posterior mean for Log Hazard is significant at the 5% level. Nevertheless, in both exercises, the evidence for the posterior mode model relative to other models in the model space was not overwhelming, and hence we continue to rely on the evidence from the BMA exercises, which, as we saw, turns out to be ambiguous across the two periods for Log Hazard.

Our analysis highlights the importance of land inequality in influencing both current institutions and schooling through the former's impact on delaying the extension of schooling

[10] 2SLS-BMA has been proposed by Durlauf, Kourtellos, and Tan (2011) in the context of just identification and extended to the case of over identification by Eicher, Lenkoski, and Raftery (2009).

opportunities. In particular, our findings agree with both Acemoglu et al and Glaeser et al. Early institutions do play a critical role in determining current institutions, but so do the initial conditions surrounding early human capital accumulation. However, at least for the sample of countries we have, we do not find evidence for an important role of early institutions in determining current human capital levels. The main explanation for the variation of current human capital levels appears to be variations in the ability of countries to substantially extend schooling opportunities early on in the development process. In turn, a key determinant of delays in reaching early schooling milestones is the inequality of land ownership.

### 3.3.2 Long-run economic performance

We now extend the analysis in the previous section to investigate the implications of schooling delays on long-run economic performance. The results from the previous section suggest that this would be accomplished through both the current schooling and current institutions channels. To facilitate our analysis, we employ the canonical cross-country income regression framework along the lines of Hall and Jones (1999), Acemoglu, Johnson, and Robinson (2001), and Glaeser, La Porta, Lopez-De-Silanes, and Shleifer (2004), which is standard in the growth literature. This regression is given by equation (3.3.3) that takes the form of a linear regression of log GDP per capita in 1995 on current institutions, $R_{T,i}$, schooling, $H_{T,i}$, and other factors, $Z_i$ that include Tropics, Ethnic Fractionalization, and colonial dummies,

$$y_{T,i} = \mu_y + \alpha_y R_{T,i} + \beta_y H_{T,i} + Z_i' \gamma_y + e_{y,i}, \qquad (3.3.3)$$

Our identification strategy exploits the variation in the cross-country distribution of land inequality and its effect on the delay of the extension of public schooling. As put forth by Engerman and Sokoloff (2002) and GMV, variations in initial climatic conditions are responsible for the distribution of land inequality. In areas where conditions are conducive for the cultivation of large scale crops, land ownership tends to be concentrated in the hands of a small group of elites. However, in areas where only small holdings are possible, land ownership tends to be more dispersed. We posit that the historical inequality of land ownership would not constitute a direct determinant of long-run income.

To the extent that we can conceive of climatic conditions as being randomly assigned to countries, we are then able, in the spirit of Acemoglu, Johnson, and Robinson (2001), to assume our key exclusion restriction. As argued in the previous section, land inequality can assert an influence over long-run income via its effect on schooling delays, which in turn affects contemporaneous determinants of current income, such as current institutions and schooling. We therefore instrument current Schooling with the Log Hazard.

However, we urge the reader to view the results in this section as being purely suggestive. As pointed out by Brock and Durlauf (2001), it is very difficult to obtain strong causal statements using cross-country growth regressions because very many factors (some of which would be invariably omitted even with a large model space) potentially determine long-run growth, and it is very difficult to argue that proposed instrumental variables are orthogonal to these factors. We therefore urge the reader to consider our findings within the environment of the existing literature, and to think of our findings as being only comparable to the existing findings in the literature that would also invariably suffer from the same criticism.

Our main strategy employs the Log Hazard and Initial Executive Constraints as instruments for current Schooling and current Executive Constraints, respectively. Table 3.7 presents our main findings for the two alternative periods: 1965-95 and 1985-95. Panel A shows BMA-2SLS results, which include the posterior inclusion probability of each variable, as well as the corresponding posterior mean and posterior standard error. We also present results for the posterior mode model as well as the largest model. Panel B presents the first-stage results.

The first stage results confirm that the Log Hazard and Initial Executive Constraints are good instruments for current Schooling and current Executive Constraints. In the first stage regression of current Executive Constraints, Initial Executive Constraints is strongly significant at the 1% level. This is true for both time periods. Similarly, the Log Hazard is also strongly significant at the 1% level for both time periods in the first stage regression for current Schooling. Finally, note that in all cases the F-statistics are well above 10 that suggest that our instruments are not weak.

In the second stage current income regression, for both time periods, we find that both current Executive Constraints and current Schooling are important determinants of long-run income. The posterior inclusion probability for both these variables are very high at close to 100%. However, while both the current Executive Constraints and current Schooling are very likely to be variables in the true model, once we account for model uncertainty, only current Schooling turns out to have a (highly) significant impact (at the 1% level) on long-run income. Our findings suggest, therefore, that land inequality has an ultimate and important influence on long-run income via the human capital channel.

Next, we provide two robustness exercises to our main strategy reported in Tables 3.8-3.9.

First, we employ the same income regression as in Table 3.7, but account for both second stage and first stage model uncertainty along the lines of Eicher, Lenkoski, and Raftery (2009). In particular, we instrument current Executive Constraints and current Schooling using the full set of historical determinants from Tables 3.5 and 3.6 to compute the 2SLS-BMA estimates. We report the results from this over-identified exercise in Table 3.8. Panel A of Table 3.8 confirms that the results we obtained for the just-identified case in Table 3.7 for Schooling are robust to the inclusion of additional instruments. As in the latter case, the posterior inclusion probabilities for Schooling are close to 100% for both time periods. The posterior means are also always strongly significant at the 1% level. However, the findings for current Executive Constraints are significantly weakened from before. Now, current Executive Constraints is found to be a far less important determinant of current income. Its posterior inclusion probability has dropped from close to 100% in Table Table 3.7 to 49.8% and 67.5% for the periods 1965-95 and 1985-95, respectively. Its posterior mean remains insignificant for both periods.

As a further robustness check, we also report results that drop current Schooling from the model space. These results are reported in Table 3.9 and Table 3.8 (Panel B), and correspond to those in Tables 3.7 and 3.8 (Panel A), respectively. This exercise provides a check that we are able to verify the existing results in the literature (e.g., Acemoglu, Johnson, and Robinson (2001) and Glaeser, La Porta, Lopez-De-Silanes, and Shleifer (2004)) that find a major role for institutional quality in determining cross-country differences in economic development. For example, Acemoglu, Johnson, and Robinson (2001) was able to do so using log settler mortality to instrument for current institutions. Unfortunately, we could not use the preferred instrument of Acemoglu, Johnson, and Robinson (2001) because it severely restricts our sample. However, when we drop current Schooling from the income regression (3.3.3), we find that current Executive Constraints is an important and highly significant determinant of long-run income. This result is consistent with the findings in the existing literature and is, therefore, precisely what we expected to find.

In terms of the other growth determinants, we find that British Colony negatively affects income at the 1% significant level and with a posterior inclusion probability of 1. This finding is consistent with the findings of Acemoglu, Johnson, and Robinson (2001). As pointed out by Acemoglu et al, a possible explanation for the negative effect of British Colony on income is that researchers are overestimating the negative quality of institutions for French colonies, and the second-stage effect of British colony is correcting for this. We also find that Ethnic Fractionalization and Tropics are not robust determinants of long-run economic performance. Interestingly, when we consider the effect of Schooling the posterior inclusion probability of Tropics is always below 50% with a negative but insignificant effect (see Tables 3.7 and 3.8 (Panel A)). However, when we do not consider the effect of Schooling (see Table 3.8 (Panel B)), Tropics appears to play an important role with posterior inclusion probability of 97.2% and 91.2% and a negative and strongly significant effect at the 1% and 5% levels, for the periods of 1965-95 and 1985-95, respectively. Our reading of this result is that Tropics is masking the correlation of geography with land inequality, which in turn is the key determinant of schooling in the first stage.

Overall, our findings highlight the important role that human capital accumulation plays in determining long-run economic performance. Since land inequality has been shown to be a key factor in determining human capital accumulation, it is therefore a crucial fundamental determinant of economic outcomes.

## 3.4 Conclusion

This paper accomplishes three things. First, we confirm the direct predictions of the theory of Galor, Moav, and Vollrath (2009) that higher levels of land inequality result in delays in the implementation of human capital enhancing policies. Using new historical data by Frankema (2009), we test the importance of land inequality as a determinant of delays in the extension of schooling opportunities against alternatives theories. Next, we examine the effect of schooling delays on contemporaneous determinants of long-run income; specifically, current institutions and human capital formation. Our findings suggest new channels through which land inequality potentially affects long-run economic performance. Finally, we contribute to the ongoing debate in the growth literature over whether it is the historical level of human capital or the historical quality of institutions that is ultimately responsible for long-run economic performance. While our findings do not allow us to assert the primacy of either of these deep determinants, they do suggest a stronger role for human capital. We certainly do not find evidence to support the hypothesis that initial institutions determine current schooling levels. Rather, our work concludes that it is land inequality and the incentives it provides to elites to delay the extension of schooling opportunities that ultimately results in the failure of countries to launch economically through the effects of schooling delays on both current schooling levels and quality of institutions.

## Table 3.1: List of Countries

This table presents the list of countries. In the bracket we note the date of the actual starting date of each country.

| **Europe** | **Latin America** | **Middle East and North Africa** |
|---|---|---|
| Austria (1919) | Argentina (1895) | Algeria (1962) |
| Belgium (1860) | Brazil (1872) | Egypt (1951) |
| Denmark (1882) | Chile (1895) | Iran (1887) |
| France (1851) | Colombia (1938) | Iraq (1957) |
| Greece (1870) | Costa Rica (1883) | Morocco (1956) |
| Ireland (1926) | Dom. Rep (1844) | Tunisia (1956) |
| Netherlands (1899) | Ecuador (1950) | |
| Norway (1855) | El Salvador (1930) | **South Asia** |
| Portugal (1864) | Guatemala (1950) | India (1947) |
| Romania (1899) | Honduras (1930) | Pakistan (1961) |
| Sweden (1882) | Nicaragua (1950) | Sri Lanka (1948) |
| UK (1860) | Panama (1950) | |
| | Paraguay (1950) | **Central Asia** |
| **Offshoots** | Peru (1961) | Turkey (1935) |
| Australia (1946) | | |
| Canada (1865) | **East Asia and Pacific** | |
| N. Zealand (1907) | China (1953) | |
| USA (1830) | Indonesia (1961) | |
| | Japan (1815) | |
| **Sub-Saharan** | Korea Rep. (1960) | |
| Ghana (1960) | Malaysia (1957) | |
| Kenya (1969) | Myanmar (1948) | |
| Mozambique (1975) | Philippines (1948) | |
| South Africa (1946) | Thailand (1929) | |
| Zambia (1969) | | |

**Table 3.2: Descriptive Statistics**

This table presents the summary statistics for the 53 countries of our dataset.

| | Mean | St. Dev. | Min | Max |
|---|---|---|---|---|
| Initial Income | 7.010 | 0.482 | 6.116 | 8.583 |
| Income in 1995 | 8.525 | 1.016 | 6.452 | 10.11 |
| Initial Schooling | 0.378 | 0.229 | 0.000 | 0.910 |
| Schooling, 1965-95 | 0.298 | 0.794 | -2.291 | 1.685 |
| Schooling, 1985-95 | 0.643 | 0.702 | -2.034 | 1.790 |
| Initial Executive Constraints | 0.370 | 0.399 | 0.000 | 1.000 |
| Executive Constraints, 1965-95 | 0.570 | 0.341 | 0.000 | 1.000 |
| Executive Constraints, 1985-95 | 0.686 | 0.309 | 0.000 | 1.000 |
| Political Instability | 0.278 | 0.385 | 0.000 | 2.375 |
| Auto Registration | 0.037 | 0.106 | 0.000 | 0.752 |
| BFLZ Gini Index | 0.491 | 0.110 | 0.276 | 0.794 |
| Land Gini | 0.636 | 0.144 | 0.307 | 0.863 |
| Arable Land | 6.934 | 1.573 | 3.367 | 11.27 |
| Independence | 84.00 | 73.10 | 0.000 | 199.0 |
| Tropics | 0.376 | 0.417 | 0.000 | 1.000 |
| Ethnic Fractionalization | 0.373 | 0.247 | 0.012 | 0.859 |
| Frecivil | 0.528 | 0.504 | 0.000 | 1.000 |
| Britcommon | 0.327 | 0.474 | 0.000 | 1.000 |
| British Colony | 0.302 | 0.463 | 0.000 | 1.000 |
| French Colony | 0.057 | 0.233 | 0.000 | 1.000 |
| Span./Port Colony | 0.302 | 0.463 | 0.000 | 1.000 |
| Other Colonies | 0.113 | 0.320 | 0.000 | 1.000 |

**Table 3.3: Primary Schooling Threshold Failure**

This table lists the countries that fail to attain the primary schooling for various thresholds.

| 40% | 45% | 50% | 55% | 60% | 65% | 70% |
|---|---|---|---|---|---|---|
| Morocco | India | Egypt | Egypt | Egypt | China | Algeria |
| Pakistan | Morocco | India | Guatemala | Ghana | Colombia | China |
| | Pakistan | Morocco | India | Guatemala | Egypt | Colombia |
| | | Mozambique | Iraq | Honduras | Ghana | Costa Rica |
| | | Pakistan | Morocco | India | Guatemala | Egypt |
| | | | Mozambique | Iran | Honduras | El Salvador |
| | | | Myanmar | Iraq | India | Ghana |
| | | | Nicaragua | Morocco | Iran | Greece |
| | | | Pakistan | Mozambique | Iraq | Guatemala |
| | | | Turkey | Myanmar | Korea Republic | Honduras |
| | | | | Nicaragua | Morocco | India |
| | | | | Pakistan | Mozambique | Iran |
| | | | | Turkey | Myanmar | Iraq |
| | | | | | Nicaragua | Japan |
| | | | | | Pakistan | Korea Republic |
| | | | | | Thailand | Morocco |
| | | | | | Turkey | Mozambique |
| | | | | | Zambia | Myanmar |
| | | | | | | Nicaragua |
| | | | | | | Pakistan |
| | | | | | | Paraguay |
| | | | | | | Thailand |
| | | | | | | Turkey |
| | | | | | | Zambia |

**Table 3.4: Hazard Model for the Delay in Primary Schooling**

The table presents BMA results for the Cox-PH duration model. The posterior inclusion probability (PIP) of a variable is the sum of the posterior probabilities models that include that variable. The posterior mean (PM) is the average of the partial likelihood coefficient estimates (COEF) of individual models weighted by posterior probability. The posterior standard error (PSE) is the BMA estimate for the standard error (SE) taking model uncertainty into account. All reported standard errors refer to heteroskedasticity consistent standard errors.

| | Model Averaging | | | Posterior Mode Models | | | | Largest Model | |
|---|---|---|---|---|---|---|---|---|---|
| | PIP | PM | PSE | COEF | SE | COEF | SE | COEF | SE |
| Initial Income | 100.0 | 1.423 | 0.466 | 1.456 | 0.390 | 1.423 | 0.400 | 1.258 | 0.719 |
| Land gini | 98.9 | -6.643 | 2.081 | -6.844 | 1.924 | -6.857 | 2.156 | -6.509 | 1.935 |
| Arable Land | 97.8 | -0.425 | 0.158 | -0.422 | 0.125 | -0.424 | 0.137 | -0.438 | 0.154 |
| British Colony | 95.5 | 2.199 | 0.896 | 2.307 | 0.613 | 2.283 | 0.621 | 2.438 | 1.151 |
| Span./Port. Colony | 9.6 | -0.048 | 0.247 | - | - | - | - | -0.265 | 0.847 |
| French Colony | 17.4 | 0.256 | 0.751 | - | - | - | - | 1.657 | 1.380 |
| Other Colonies | 100.0 | 2.687 | 0.785 | 2.623 | 0.644 | 2.662 | 0.630 | 3.288 | 0.899 |
| Independence | 100.0 | -0.024 | 0.005 | -0.024 | 0.005 | -0.023 | 0.004 | -0.028 | 0.007 |
| Political Instability | 100.0 | -2.392 | 0.886 | -2.313 | 0.741 | -2.576 | 0.823 | -2.426 | 0.947 |
| Initial Executive Constraints | 15.2 | 0.141 | 0.479 | - | - | - | - | 0.763 | 0.779 |
| Tropics | 13.6 | -0.07 | 0.262 | - | - | - | - | 0.148 | 0.853 |
| Ethnic Fractionalization | 37.6 | -0.519 | 0.817 | - | - | -1.411 | 0.698 | -1.170 | 0.879 |
| Auto Registration | 6.4 | -0.012 | 0.438 | - | - | - | - | 0.410 | 1.550 |
| BFLZ Gini Index | 11.9 | 0.199 | 0.832 | - | - | - | - | 2.385 | 2.346 |
| Wald statistic | | | | 42.79 | | 41.02 | | 63.38 | |
| Posterior Model Probabality | | | | 0.194 | | 0.175 | | 0.000 | |

## Table 3.5: Historical Determinants of Current Schooling

The table presents 2SLS-BMA results for average schooling for the periods 1965-95 and 1985-95. The posterior inclusion probability (PIP) of a variable is the sum of the posterior probabilities models that include that variable. The posterior mean (PM) is the average of the coefficient estimates (COEF) of individual models weighted by posterior probability. The posterior standard error (PSE) is the BMA estimate for the standard error (SE) taking model uncertainty into account. All reported standard errors refer to heteroskedasticity consistent standard errors.

*Panel A: Schooling in 1965-1995*

| | Model Averaging | | | Posterior Mode | | Largest | |
|---|---|---|---|---|---|---|---|
| | PIP | PM | PSE | COEF | SE | COEF | SE |
| Constant | 100.0 | 0.470 | 0.361 | 0.493 | 0.331 | 0.606 | 0.294 |
| British Colony | 16.0 | -0.004 | 0.175 | - | - | 0.272 | 0.471 |
| Span./Port Colony | 21.6 | 0.042 | 0.237 | - | - | 0.216 | 0.374 |
| French Colony | 27.3 | -0.135 | 0.519 | - | - | -0.423 | 0.758 |
| Other Colonies | 26.2 | 0.087 | 0.256 | - | - | 0.297 | 0.373 |
| Initial Exec. Constraints | 17.6 | 0.034 | 0.212 | - | - | 0.26 | 0.504 |
| Initial Schooling | 31.6 | 0.150 | 0.385 | - | - | 0.366 | 0.547 |
| Log Hazard | 100.0 | 0.207 | 0.101 | 0.223 | 0.102 | 0.161 | 0.093 |
| Frecivil | 81.7 | -0.357 | 0.360 | -0.377 | 0.328 | -0.543 | 0.336 |
| Britcommon | 19.5 | -0.043 | 0.218 | - | - | -0.538 | 0.577 |
| Tropics | 97.8 | -0.614 | 0.364 | -0.591 | 0.377 | -0.694 | 0.284 |
| Ethnic Fractionalization | 22.9 | -0.088 | 0.335 | - | - | -0.382 | 0.569 |

*Panel B: Schooling in 1985-1995*

| | Model Averaging | | | Posterior Mode | | Largest | |
|---|---|---|---|---|---|---|---|
| | PIP | PM | PSE | COEF | SE | COEF | SE |
| Constant | 100.0 | 0.746 | 0.312 | 0.810 | 0.288 | 0.877 | 0.273 |
| British Colony | 15.44 | 0.005 | 0.147 | - | - | 0.245 | 0.44 |
| Span./Port Colony | 18.36 | -0.014 | 0.169 | - | - | 0.106 | 0.347 |
| French Colony | 14.43 | -0.026 | 0.317 | - | - | -0.127 | 0.699 |
| Other Colonies | 25.41 | 0.069 | 0.203 | - | - | 0.283 | 0.331 |
| Initial Exec. Constraints | 16.6 | 0.023 | 0.165 | - | - | 0.169 | 0.437 |
| Initial Schooling | 28.19 | 0.114 | 0.318 | - | - | 0.284 | 0.499 |
| Log Hazard | 100 | 0.185 | 0.09 | 0.193 | 0.090 | 0.149 | 0.088 |
| Frecivil | 62.9 | -0.201 | 0.275 | -0.287 | 0.278 | -0.399 | 0.297 |
| Britcommon | 16.73 | -0.018 | 0.16 | - | - | -0.412 | 0.493 |
| Tropics | 97.3 | -0.549 | 0.315 | -0.558 | 0.314 | -0.579 | 0.262 |
| Ethnic Fractionalization | 21.63 | -0.073 | 0.29 | - | - | -0.329 | 0.523 |

# Table 3.6: Historical Determinants of Current Institutions

The table presents 2SLS-BMA results for average executive constraints for the periods 1965-95 and 1985-95. The posterior inclusion probability (PIP) of a variable is the sum of the posterior probabilities models that include that variable. The posterior mean (PM) is the average of the coefficient estimates (COEF) of individual models weighted by posterior probability. The posterior standard error (PSE) is the BMA estimate for the standard error (SE) taking model uncertainty into account. All reported standard errors refer to heteroskedasticity consistent standard errors.

*Panel A: Executive Constraints in 1965-1995*

| | Model Averaging | | | Posterior Mode | | Largest | |
|---|---|---|---|---|---|---|---|
| | PIP | PM | PSE | COEF | SE | COEF | SE |
| Constant | 100.0 | 0.500 | 0.153 | 0.511 | 0.110 | 0.598 | 0.175 |
| British Colony | 33.8 | -0.075 | 0.163 | - | - | -0.080 | 0.187 |
| Span./Port Colony | 17.4 | 0.003 | 0.076 | - | - | -0.009 | 0.181 |
| French Colony | 65.7 | -0.265 | 0.315 | -0.430 | 0.295 | -0.441 | 0.272 |
| Other Colonies | 30.6 | -0.051 | 0.119 | - | - | -0.182 | 0.199 |
| Initial Exec. Constraints | 98.9 | 0.504 | 0.249 | 0.568 | 0.232 | 0.599 | 0.196 |
| Initial Schooling | 13.3 | 0.006 | 0.096 | - | - | 0.079 | 0.226 |
| Log Hazard | 99.3 | 0.071 | 0.035 | 0.064 | 0.031 | 0.055 | 0.031 |
| Frecivil | 30.5 | -0.041 | 0.111 | - | - | -0.110 | 0.185 |
| Britcommon | 56.0 | -0.186 | 0.243 | -0.305 | 0.192 | -0.368 | 0.236 |
| Tropics | 68.9 | -0.140 | 0.154 | -0.227 | 0.136 | -0.200 | 0.137 |
| Ethnic Fractionalization | 15.5 | -0.015 | 0.114 | - | - | -0.014 | 0.241 |

*Panel B: Executive Constraints in 1985-1995*

| | Model Averaging | | | Posterior Mode | | Largest | |
|---|---|---|---|---|---|---|---|
| | PIP | PM | PSE | COEF | SE | COEF | SE |
| Constant | 100.0 | 0.691 | 0.115 | 0.694 | 0.087 | 0.753 | 0.138 |
| British Colony | 31.1 | -0.086 | 0.174 | - | - | -0.066 | 0.164 |
| Span./Port Colony | 15.5 | 0.007 | 0.052 | - | - | 0.069 | 0.139 |
| French Colony | 90.3 | -0.448 | 0.263 | -0.504 | 0.226 | -0.479 | 0.217 |
| Other Colonies | 17.5 | -0.016 | 0.068 | - | - | -0.048 | 0.157 |
| Initial Exec. Constraints | 98.7 | 0.530 | 0.206 | 0.557 | 0.187 | 0.549 | 0.160 |
| Initial Schooling | 13.9 | 0.008 | 0.074 | - | - | 0.033 | 0.178 |
| Log Hazard | 89.7 | 0.044 | 0.027 | 0.046 | 0.021 | 0.039 | 0.021 |
| Frecivil | 17.7 | -0.013 | 0.064 | - | - | -0.087 | 0.155 |
| Britcommon | 77.0 | -0.342 | 0.248 | -0.450 | 0.159 | -0.412 | 0.199 |
| Tropics | 78.4 | -0.164 | 0.126 | -0.210 | 0.100 | -0.210 | 0.092 |
| Ethnic Fractionalization | 19.5 | -0.030 | 0.113 | - | - | -0.069 | 0.201 |

# Table 3.7: Long-run Income Regression: including schooling (just identification)

The table presents 2SLS-BMA results for log per capita income in 1995 using fitted Log Hazard and Initial Executive Constraints as instruments for current schooling and executive constraints, respectively. The posterior inclusion probability (PIP) of a variable is the sum of the posterior probabilities models that include that variable. The posterior mean (PM) is the average of the coefficient estimates (COEF) of individual models weighted by posterior probability. The posterior standard error (PSE) is the BMA estimate for the standard error (SE) taking model uncertainty into account. All reported standard errors refer to heteroskedasticity consistent standard errors. Panel B presents the first stage results for both Executive Constraints and Schooling for 1965-95 and 1985-1995.

*Panel A: Second stage results*

| | Model Averaging | | | Posterior Mode | | Largest | | Model Averaging | | | Posterior Mode | | Largest | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | PIP | PM | PSE | COEF | SE | COEF | SE | PIP | PM | PSE | COEF | SE | COEF | SE |
| Constant | 100.0 | 8.031 | 0.269 | 8.078 | 0.249 | 7.815 | 0.304 | 100.0 | 7.292 | 0.290 | 7.271 | 0.294 | 7.029 | 0.342 |
| British Colony | 100.0 | -0.681 | 0.132 | -0.679 | 0.118 | -0.770 | 0.215 | 100.0 | -0.539 | 0.155 | -0.531 | 0.149 | -0.678 | 0.212 |
| Span./Port Colony | 14.2 | -0.004 | 0.076 | - | - | 0.033 | 0.226 | 13.1 | -0.001 | 0.085 | - | - | -0.014 | 0.212 |
| French Colony | 8.9 | 0.041 | 0.146 | - | - | 0.547 | 0.280 | 8.4 | 0.024 | 0.109 | - | - | 0.355 | 0.260 |
| Other Colonies | 14.8 | -0.016 | 0.113 | - | - | -0.112 | 0.300 | 17.7 | -0.042 | 0.159 | - | - | -0.272 | 0.269 |
| Exec. Con., 1965-95 | 99.8 | 0.809 | 0.559 | 0.829 | 0.552 | 0.801 | 0.540 | - | - | - | - | - | | |
| Schooling, 1965-95 | 100.0 | 0.983 | 0.287 | 0.945 | 0.298 | 1.117 | 0.341 | - | - | - | - | - | | |
| Exec. Con., 1985-95 | - | - | - | - | - | - | - | 99.7 | 1.032 | 0.607 | 1.010 | 0.606 | 1.103 | 0.596 |
| Schooling, 1985-95 | - | - | - | - | - | - | - | 99.7 | 1.128 | 0.346 | 1.146 | 0.318 | 1.246 | 0.400 |
| Tropics | 48.9 | -0.102 | 0.198 | -0.212 | 0.232 | -0.163 | 0.288 | 35.3 | -0.061 | 0.179 | - | - | -0.128 | 0.284 |
| Ethnic Fractionalization | 4.7 | 0.018 | 0.116 | - | - | 0.583 | 0.359 | 3.5 | 0.015 | 0.110 | - | - | 0.609 | 0.384 |

*Panel B: First stage results*

| | 1965-95 | | | | 1985-95 | | | |
|---|---|---|---|---|---|---|---|---|
| | Exec. Con. | | Schooling | | Exec. Con. | | Schooling | |
| | COEF | SE | COEF | SE | COEF | SE | COEF | SE |
| Constant | 0.552 | 0.082 | 0.443 | 0.158 | 0.704 | 0.087 | 0.761 | 0.139 |
| British Colony | -0.312 | 0.139 | 0.043 | 0.262 | -0.343 | 0.145 | 0.068 | 0.211 |
| Span./Port Colony | -0.047 | 0.094 | -0.089 | 0.176 | 0.048 | 0.080 | -0.115 | 0.148 |
| French Colony | -0.462 | 0.089 | -0.633 | 0.190 | -0.490 | 0.108 | -0.277 | 0.160 |
| Other Colonies | -0.181 | 0.089 | 0.149 | 0.230 | -0.039 | 0.099 | 0.180 | 0.205 |
| Initial Exec. Con. | 0.526 | 0.127 | 0.250 | 0.252 | 0.459 | 0.135 | 0.157 | 0.200 |
| Log Hazard | 0.063 | 0.016 | 0.193 | 0.043 | 0.045 | 0.017 | 0.174 | 0.042 |
| Tropics | -0.180 | 0.087 | -0.594 | 0.170 | -0.192 | 0.078 | -0.506 | 0.146 |
| Ethnic Fractionalization | -0.038 | 0.162 | -0.498 | 0.279 | -0.088 | 0.150 | -0.414 | 0.269 |
| F-stat | 35.41 | | 25.15 | | 17.01 | | 28.67 | |

## Table 3.8: Long-run Income Regression (over identification)

The table presents 2SLS-BMA results for log per capita income in 1995 using an over-identification strategy based on the results in Tables 3.6 and 3.5 as the first stage. Panel A presents the results when we account for the effect of both current schooling and executive constraints while Panel B presents the results for the case when we exclude the effect of current schooling. The posterior inclusion probability (PIP) of a variable is the sum of the posterior probabilities models that include that variable. The posterior mean (PM) is the average of the coefficient estimates (COEF) of individual models weighted by posterior probability. The posterior standard error (PSE) is the BMA estimate for the standard error (SE) taking model uncertainty into account. All reported standard errors refer to heteroskedasticity consistent standard errors.

*Panel A: Including Schooling*

| | Model Averaging | | | Posterior Mode | | Largest | | Model Averaging | | | Posterior Mode | | Largest | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | PIP | PM | PSE | COEF | SE | COEF | SE | PIP | PM | PSE | COEF | SE | COEF | SE |
| Constant | 100.0 | 8.107 | 0.321 | 8.362 | 0.117 | 7.829 | 0.243 | 100.0 | 7.331 | 0.374 | 7.160 | 0.218 | 7.061 | 0.280 |
| British Colony | 99.5 | -0.693 | 0.161 | -0.660 | 0.148 | -0.742 | 0.199 | 96.0 | -0.530 | 0.193 | -0.527 | 0.140 | -0.549 | 0.203 |
| Span./Port Colony | 14.0 | 0.011 | 0.091 | - | - | 0.064 | 0.226 | 14.6 | 0.014 | 0.100 | - | - | 0.010 | 0.218 |
| French Colony | 13.0 | 0.003 | 0.124 | - | - | 0.091 | 0.312 | 15.5 | 0.014 | 0.159 | - | - | 0.219 | 0.243 |
| Other Colonies | 12.5 | -0.003 | 0.103 | - | - | 0.027 | 0.252 | 13.5 | -0.013 | 0.123 | - | - | -0.106 | 0.293 |
| Exec. Con., 1965-95 | 49.8 | 0.612 | 0.742 | - | - | 1.315 | 0.418 | - | - | - | - | - | - | - |
| Schooling, 1965-95 | 97.1 | 1.087 | 0.338 | 1.321 | 0.129 | 0.791 | 0.256 | - | - | - | - | - | - | - |
| Exec. Con., 1985-95 | - | - | - | - | - | - | - | 67.5 | 0.815 | 0.675 | 1.134 | 0.402 | 1.408 | 0.478 |
| Schooling, 1985-95 | - | - | - | - | - | - | - | 99.9 | 1.269 | 0.304 | 1.181 | 0.236 | 1.060 | 0.314 |
| Tropics | 24.8 | -0.083 | 0.203 | - | - | -0.364 | 0.263 | 16.6 | -0.032 | 0.132 | - | - | -0.198 | 0.284 |
| Ethnic Fractionalization | 13.2 | 0.018 | 0.137 | - | - | 0.174 | 0.335 | 13.3 | 0.015 | 0.148 | - | - | 0.185 | 0.371 |

*Panel B: Excluding Schooling*

| | Model Averaging | | | Posterior Mode | | Largest | | Model Averaging | | | Posterior Mode | | Largest | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | PIP | PM | PSE | COEF | SE | COEF | SE | PIP | PM | PSE | COEF | SE | COEF | SE |
| Constant | 100.0 | 7.505 | 0.254 | 7.239 | 0.338 | 7.566 | 0.362 | 100.0 | 6.862 | 0.498 | 6.799 | 0.430 | 6.925 | 0.514 |
| British Colony | 99.9 | -0.822 | 0.201 | -0.694 | 0.232 | -0.860 | 0.268 | 60.1 | -0.254 | 0.275 | -0.199 | 0.238 | -0.484 | 0.266 |
| Span./Port Colony | 16.1 | -0.026 | 0.118 | - | - | -0.139 | 0.269 | 19.0 | -0.041 | 0.156 | - | - | -0.279 | 0.254 |
| French Colony | 12.8 | 0.014 | 0.121 | - | - | 0.029 | 0.327 | 28.6 | 0.228 | 0.437 | - | - | 0.448 | 0.387 |
| Other Colonies | 12.6 | 0.005 | 0.116 | - | - | -0.020 | 0.350 | 13.4 | -0.009 | 0.161 | - | - | -0.212 | 0.407 |
| Exec. Con., 1965-95 | 100.0 | 2.723 | 0.331 | 2.980 | 0.378 | 2.704 | 0.420 | - | - | - | - | - | - | - |
| Exec. Con., 1985-95 | - | - | - | - | - | - | - | 100.0 | 2.981 | 0.508 | 3.053 | 0.456 | 3.080 | 0.550 |
| Tropics | 97.2 | -0.703 | 0.264 | - | - | -0.646 | 0.286 | 91.2 | -0.681 | 0.334 | - | - | -0.542 | 0.318 |
| Ethnic Fractionalization | 13.0 | -0.016 | 0.169 | -0.568 | 0.433 | -0.078 | 0.464 | 24.4 | -0.128 | 0.327 | -0.869 | 0.437 | -0.204 | 0.475 |

## Table 3.9: Long-run Income Regression: excluding schooling (just identification)

The table presents 2SLS-BMA results for log per capita income in 1995 using initial institutions as instrument for current institutions. The posterior inclusion probability (PIP) of a variable is the sum of the posterior probabilities models that include that variable. The posterior mean (PM) is the average of the coefficient estimates (COEF) of individual models weighted by posterior probability. The posterior standard error (PSE) is the BMA estimate for the standard error (SE) taking model uncertainty into account. All reported standard errors refer to heteroskedasticity consistent standard errors. Panel B presents the first stage results for Executive constraints for 1965-95 and 1985-1995.

*Panel A: Second stage results*

| | Model Averaging | | | Posterior Mode | | Largest | | Model Averaging | | | Posterior Mode | | Largest | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | PIP | PM | PSE | COEF | SE | COEF | SE | PIP | PM | PSE | COEF | SE | COEF | SE |
| Constant | 100.0 | 8.038 | 0.260 | 8.002 | 0.241 | 8.194 | 0.331 | 100.0 | 7.600 | 0.370 | 7.552 | 0.351 | 7.750 | 0.450 |
| British Colony | 100.0 | -0.704 | 0.189 | -0.717 | 0.176 | -0.636 | 0.244 | 99.9 | -0.460 | 0.230 | -0.448 | 0.211 | -0.500 | 0.240 |
| Span./Port Colony | 16.8 | -0.031 | 0.112 | - | - | -0.087 | 0.245 | 25.7 | -0.080 | 0.180 | - | - | -0.280 | 0.230 |
| French Colony | 12.8 | -0.016 | 0.107 | - | - | -0.191 | 0.291 | 12.3 | 0.010 | 0.120 | - | - | -0.040 | 0.340 |
| Other Colonies | 20.0 | 0.058 | 0.157 | - | - | 0.243 | 0.269 | 14.6 | 0.020 | 0.130 | - | - | 0.010 | 0.320 |
| Exec. Con., 1965-95 | 100.0 | 1.865 | 0.339 | 1.892 | 0.333 | 1.752 | 0.353 | - | - | - | - | - | - | - |
| Exec. Con., 1985-95 | - | - | - | - | - | - | - | 100.0 | 2.110 | 0.430 | 2.132 | 0.416 | 2.070 | 0.450 |
| Tropics | 98.9 | -0.820 | 0.237 | -0.852 | 0.204 | -0.763 | 0.272 | 98.2 | -0.840 | 0.260 | -0.907 | 0.202 | -0.690 | 0.290 |
| Ethnic Fractionalization | 20.3 | -0.085 | 0.237 | - | - | -0.427 | 0.363 | 21.8 | -0.090 | 0.260 | - | - | -0.370 | 0.400 |

*Panel B: First stage results*

| | 1965-95 | | 1985-95 | |
|---|---|---|---|---|
| | Exec.Con. | | Exec.Con. | |
| | COEF | SE | COEF | SE |
| Constant | 0.718 | 0.082 | 0.820 | 0.070 |
| British Colony | -0.397 | 0.163 | -0.400 | 0.160 |
| Span./Port Colony | -0.073 | 0.105 | 0.030 | 0.090 |
| French Colony | -0.652 | 0.081 | -0.620 | 0.090 |
| Other Colonies | -0.174 | 0.112 | -0.030 | 0.110 |
| Initial Exec. Con. | 0.616 | 0.144 | 0.520 | 0.140 |
| Tropics | -0.223 | 0.100 | -0.220 | 0.080 |
| Ethnic Fractionalization | -0.237 | 0.166 | -0.230 | 0.150 |
| F-stat | 19.45 | | 13.86 | |

## Table A1: Data Appendix

| Variable | Description |
|---|---|
| Delay in Schooling | Following the methodology of Comin and Hobijn (2004), we construct historical data for primary schooling enrollments, measured as the number of students in primary school as a fraction of population between 5-14. First, we verify the dataset of Comin and Hobijn (2004), which is limited to 23 industrialized countries and then expand it to 53 countries. Using this new dataset we create the delay in schooling variable, which is the time the time it takes for each country to first reach a threshold level in primary schooling enrollment, minus the time it took the first country to pass that threshold. Source: Mitchell (1998) for the population data; Banks (1999) for the number of students. |
| Initial Income | Log of GDP per capita, where for the colonies we use the independence date or earliest available, and for the non-colonies the average of earliest possible until 1831 (threshold for 50%). Source: Maddison (2009). |
| Income in 1995 | Log of GDP per capita 1995. Source: Maddison (2009). |
| Initial Schooling | Primary schooling enrollments is based on authors' calculations using historical schooling data. For colonies we use the independence date (or earliest available if the independence not available). For non-colonies, we use the earliest available. Source: Mitchell (1998), Banks (1999). |
| Schooling | Logarithm of average years of male secondary and higher school attainment (25+), average for the periods 1965- 1995 and 1985-1995. Source: Barro and Lee (2010). |
| Initial Executive Constraints | Institutional variable with the lowest value 0 indicating unlimited executive authority and 1 executive parity or subordination. For the colonies we use the independence date or earliest available, and for the non-colonies the average of earliest possible until 1831 (threshold for 50%). Source: Polity IV. |
| Executive constraints | Institutional variable with the lowest value 0 indicating unlimited executive authority and 1 executive parity or subordination, average for the periods 1965- 1995 and 1985-1995. Source: Polity IV. |
| Auto Registration | Number of passenger cars (excluding tractors and similar vehicles) in use. Numbers typically derived from registration and licensing records, meaning that vehicles out of use may occasionally be included. We divide the variable to population and then multiply by 10. We use data over the period 1895-1978. Source: Comin and Hobijn (2004). |
| BFLZ Gini Index | BFLZ Gini Index is based on a large number of observations of within country inequality spanning from 1820-1995. For non-colonies we use earliest available and for colonies we use the Independence date and if not available we use the earliest available after Independence. In particular, the Gini Index is based on direct income Gini estimates; estimates of the net household or expenditure Ginis; Ginis based on income shares; Williamson index, which is the ratio between GDP per capita and real wages of unskilled laborers; and height inequality data. Source: Van Zanden, Baten, Földvari, and Van Leeuwen (2011). |

Table A1 continued

| **Variable** | **Description** |
|---|---|
| Land Gini | The gini coefficient of the size distribution of land. For all countries we use the earliest observation available. We use data over the period 1880-1999. Source: Frankema (2009). |
| Arable Land | Log of arable land (absolute) in hectares, in 1700. Source: Ramankutty and Foley (1999) |
| Independence | Independence The time it takes for each country to declare independence relative to the United States who declared independence in 1776. Source: CIA Factbook. |
| Tropics | Percentage of land area classified as tropical and subtropical via the Koeppen-Geiger system. Source: CID at Harvard. |
| Ethnic Fractionalization | Variable which combines racial and linguistic characteristics. Source: Alesina, Devleeschauwer, Easterly, Kurlat, and Wacziarg (2003). |
| Colonial Dummies | Coded zero or one. One indicates that country was colonized by Britain, France, Spain or Portugal. Source: CIA Factbook |
| Political Instability | Political Instability is measured as the average of the first differences (in absolute values) of the Polity2 variable from Polity IV. The Polity2 variable is a measure of the degree of democracy in a country with a score of +10 representing most democratic and -10 signifying most autocratic. The averages of the first differences are calculated as follows: for colonies we average values of the (absolute) year-to-year changes in the Polity2 variable from the year of independence to the year the colony achieves the schooling threshold, while for non-colonies, we take the corresponding average values from the earliest available observation until the year the country achieves the schooling threshold. Source: Polity IV. |
| Frecivil | Coded zero or one. It indicates that a country was colonized by France, Spain, Belgium, Portugal or Germany and French legal code was transferred. Source: La Porta, Lopez-de Silanes, Shleifer, and Vishny (1999). |
| Britcommon | Coded zero or one. It indicates that a country was colonized by Britain and English legal code was transferred. Source: La Porta, Lopez-de Silanes, Shleifer, and Vishny (1999). |

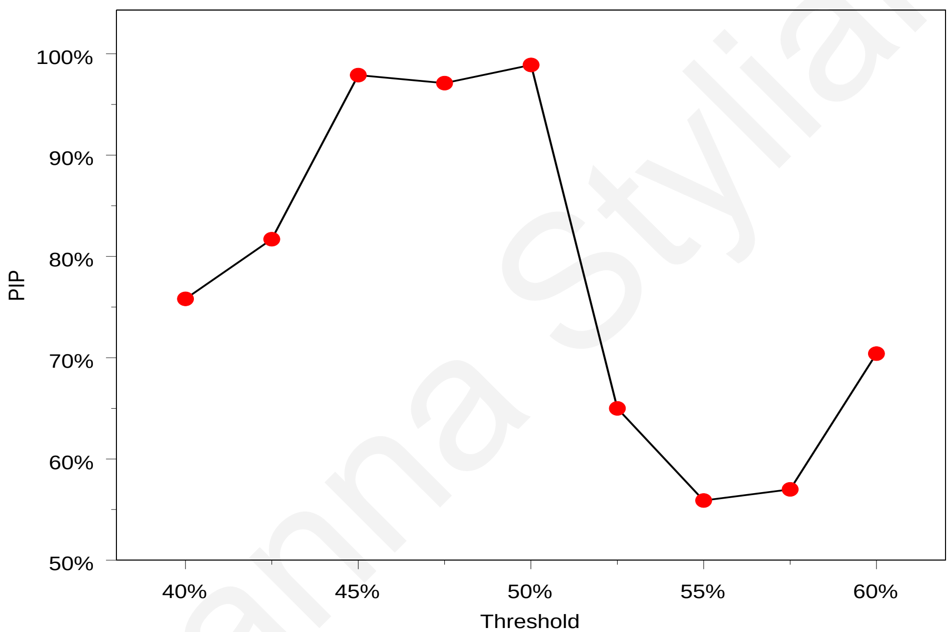
# Table A2: Hazard Model for the Delay in Primary Schooling

The table presents BMA results for the Cox-PH duration model for four different model spaces. The posterior inclusion probability (PIP) of a variable is the sum of the posterior probabilities models that include that variable. The posterior mean (PM) is the average of the partial likelihood coefficient estimates (COEF) of individual models weighted by posterior probability. The posterior standard error (PSE) is the BMA estimate for the standard error (SE) taking model uncertainty into account. All reported standard errors refer to heteroskedasticity consistent standard errors.

| | Model Averaging I | | | | | | | Model Averaging II | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | BMA | | | Posterior Mode Models | | | | BMA | | | Posterior Mode Models | | | |
| | PIP | PM | PSE | COEF | SE | COEF | SE | PIP | PM | PSE | COEF | SE | COEF | SE |
| Initial Income | 100 | 1.424 | 0.461 | 1.456 | 0.390 | 1.423 | 0.400 | 100 | 1.423 | 0.466 | 1.456 | 0.390 | 1.423 | 0.400 |
| Land gini | 98.6 | -6.58 | 2.112 | -6.844 | 1.924 | -6.857 | 2.156 | 98.9 | -6.643 | 2.081 | -6.844 | 1.924 | -6.857 | 2.156 |
| Arable Land | 97.3 | -0.423 | 0.161 | -0.422 | 0.125 | -0.424 | 0.137 | 97.8 | -0.425 | 0.158 | -0.422 | 0.125 | -0.424 | 0.137 |
| British Colony | 94.5 | 2.156 | 0.918 | 2.307 | 0.613 | 2.283 | 0.621 | 95.5 | 2.199 | 0.896 | 2.307 | 0.613 | 2.283 | 0.621 |
| Span./Port. Colony | 11.8 | -0.059 | 0.273 | - | - | - | - | 9.6 | -0.048 | 0.247 | - | - | - | - |
| French Colony | 17.6 | 0.248 | 0.737 | - | - | - | - | 17.4 | 0.256 | 0.751 | - | - | - | - |
| Other Colonies | 100 | 2.67 | 0.784 | 2.623 | 0.644 | 2.662 | 0.630 | 100 | 2.687 | 0.785 | 2.623 | 0.644 | 2.662 | 0.630 |
| Independence | 100 | -0.023 | 0.005 | -0.024 | 0.005 | -0.023 | 0.004 | 100 | -0.024 | 0.005 | -0.024 | 0.005 | -0.023 | 0.004 |
| Political Instability | 100 | -2.393 | 0.884 | -2.313 | 0.741 | -2.576 | 0.823 | 100 | -2.392 | 0.886 | -2.313 | 0.741 | -2.576 | 0.823 |
| Initial Executive Constraints | 17.3 | 0.163 | 0.515 | - | - | - | - | 15.2 | 0.141 | 0.479 | - | - | - | - |
| Tropics | 14.8 | -0.073 | 0.269 | - | - | - | - | 13.6 | -0.07 | 0.262 | - | - | - | - |
| Ethnic Fractionalization | 39 | -0.538 | 0.826 | | | -1.411 | 0.698 | 37.6 | -0.519 | 0.817 | | | -1.411 | 0.698 |
| Auto Registration | - | - | - | - | - | - | - | 6.4 | -0.012 | 0.438 | - | - | - | - |
| BFLZ Gini Index | - | - | - | - | - | - | - | 11.9 | 0.199 | 0.832 | - | - | - | - |
| Wald Statistic | | | | 42.79 | | 41.02 | | | | | 42.79 | | 41.02 | |
| Posterior Model Probability | | | | 0.237 | | 0.214 | | | | | 0.194 | | 0.175 | |

| | Model Averaging III | | | | | | | Model Averaging IV | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | BMA | | | Posterior Mode Models | | | | BMA | | | Posterior Mode Models | | | |
| | PIP | PM | PSE | COEF | SE | COEF | SE | PIP | PM | PSE | COEF | SE | COEF | SE |
| Initial Income | 100 | 1.42 | 0.461 | 1.456 | 0.390 | 1.423 | 0.400 | 100.0 | 1.430 | 0.470 | 1.456 | 0.390 | 1.423 | 0.400 |
| Land gini | 98.8 | -6.633 | 2.091 | -6.844 | 1.924 | -6.857 | 2.156 | 98.7 | -6.600 | 2.100 | -6.844 | 1.924 | -6.857 | 2.156 |
| Arable Land | 97.6 | -0.425 | 0.159 | -0.422 | 0.125 | -0.424 | 0.137 | 97.5 | -0.420 | 0.160 | -0.422 | 0.125 | -0.424 | 0.137 |
| British Colony | 95.2 | 2.185 | 0.905 | 2.307 | 0.613 | 2.283 | 0.621 | 94.9 | 2.170 | 0.910 | 2.307 | 0.613 | 2.283 | 0.621 |
| Span./Port. Colony | 10.3 | -0.051 | 0.255 | - | - | - | - | 10.9 | -0.050 | 0.260 | - | - | - | - |
| French Colony | 17.5 | 0.256 | 0.753 | - | - | - | - | 17.5 | 0.250 | 0.740 | - | - | - | - |
| Other Colonies | 100 | 2.684 | 0.786 | 2.623 | 0.644 | 2.662 | 0.630 | 100.0 | 2.670 | 0.780 | 2.623 | 0.644 | 2.662 | 0.630 |
| Independence | 100 | -0.024 | 0.005 | -0.024 | 0.005 | -0.023 | 0.004 | 100.0 | -0.020 | 0.010 | -0.024 | 0.005 | -0.023 | 0.004 |
| Political Instability | 100 | -2.391 | 0.886 | -2.313 | 0.741 | -2.576 | 0.823 | 100.0 | -2.390 | 0.880 | -2.313 | 0.741 | -2.576 | 0.823 |
| Initial Executive Constraints | 16.2 | 0.151 | 0.494 | - | - | - | - | 16.1 | 0.150 | 0.500 | - | - | - | - |
| Tropics | 14.5 | -0.074 | 0.27 | - | - | - | - | 13.7 | -0.070 | 0.260 | - | - | - | - |
| Ethnic Fractionalization | 37.4 | -0.516 | 0.815 | - | - | -1.411 | 0.698 | 39.0 | -0.540 | 0.830 | - | - | -1.411 | 0.698 |
| Auto Registration | - | - | - | - | - | - | - | 7.2 | -0.010 | 0.470 | - | - | - | - |
| BFLZ Gini Index | 12.7 | 0.212 | 0.858 | - | - | - | - | - | - | - | - | - | - | - |
| Wald Statistic | | | | 42.79 | | 41.02 | | | | | 42.79 | | 41.02 | |
| Posterior Model Probability | | | | 0.207 | | 0.187 | | | | | 0.220 | | 0.199 | |

Figure A1: Posterior inclusion probabilities for various schooling thresholds



# Chapter 4

# Robust Multiple Regimes in Volatility

## 4.1 Introduction

The seminal work of Ramey and Ramey (1995) on the adverse effects of volatility on economic growth has led to a considerable amount of interest in the need to understand the sources of growth volatility. Notable examples include Acemoglu, Johnson, Robinson, and Thaicharoend (2003) who find that institutions are a fundamental determinant of volatility through a number of microeconomic and macroeconomic channels, Mobarak (2005) who emphasizes the role of democracy in reducing growth instability, and Bekaert, Harvey, and Lundblad (2006) who provide strong evidence that financial liberalization is associated with lower growth volatility. More recently, Malik and Temple (2009) find an especially important role for geography since remote countries are more likely to have undiversified exports.

Despite all this work there is remarkably little consensus on which determinants are the most important sources of growth volatility. We posit that the major reason for this problem is that the existing empirical studies generally ignore model uncertainty that typically plague cross-country regressions. Therefore, the objective of this paper is to identify robust determinants of growth volatility using an econometric methodology that allows us to deal with two key facets of model uncertainty: theory uncertainty and parameter heterogeneity.

The term theory uncertainty was first coined by Brock and Durlauf (2001) to refer to the idea that new growth theories are open-ended, which means that any given theory of growth does not logically exclude other theories from also being relevant. In the present context, theory uncertainty implies that in the empirical modeling of growth volatility there is no a priori justification for focusing on a specific subset of explanatory variables. For example while Di Giovanni and Levchenko (2009) emphasize the importance of trade openness as a growth volatility determinant, Acemoglu, Johnson, Robinson, and Thaicharoend (2003) argue that institutions is the main source of growth volatility. It is not clear if the correct model specification should include both theories, or just one (or none) of them, since the inclusion of one theory; e.g., trade openness, does not automatically preclude the other; e.g., institutions, from also being a determinant of growth volatility. However, the estimated partial effect, say, of any particular determinant on growth volatility may vary dramatically across model specifications. How should one deal with the dependence of inference on model specifications?

One way to deal with this problem is to employ Bayesian Model Averaging (BMA), which dates back to Leamer (1978), and was further studied by Draper (1995), Kass and Raftery

(1995), and Raftery, Madigan, and Hoeting (1997). Model averaging constructs estimates that do not depend on a particular model specification but rather use information from all candidate models. In particular, it amounts to forming a weighted average of model specific estimates where the weights are given by the posterior model probabilities. BMA has been widely applied in growth regressions and has proven to be particularly useful in identifying robust growth determinants; see for example, Brock and Durlauf (2001), Fernandez, Ley, and Steel (2001), Sala-I-Martin, X. and Doppelhofer, G. and Miller, R. (2004), Durlauf, Kourtellos, and Tan (2008), and Masanjala and Papageorgiou (2008). However, in the context of growth volatility the benefits of BMA have been largely ignored. A notable exception is the paper by Malik and Temple (2009) who employ BMA to identify structural determinants of output volatility in developing countries.

The second source of model uncertainty is parameter heterogeneity, which refers to the idea that the data generating process that describes the stochastic phenomenon of growth volatility is not common for all observations (countries). There are reasons to believe that different countries may follow different growth volatility processes. For example, countries that are facing structural adjustment issues; such as those experiencing particularly high debt-to-GDP ratios or hyper-inflation, may face greater financing constraints that reduce the ability of countries to smooth out income across time. Alternatively, policy instruments that aim to stabilize growth may have different effects for countries at different levels of development. While the issue of parameter heterogeneity in growth regressions has been investigated thoroughly by a number of papers, this issue has not been systematically addressed in growth volatility regressions.

One approach that deals with the problem of parameter heterogeneity is to use threshold regression models or classification algorithms such as a regression tree. These models classify observations into stochastic processes depending on whether the observed value of a threshold variable is above (or below) a threshold value. In a seminal paper, Durlauf and Johnson (1995) employed a regression tree approach to uncover multiple growth regimes in the data. Following a similar strategy Papageorgiou (2002) organized countries into multiple growth regimes using the trade share and Tan (2010) classified countries into development clubs using the average expropriation risk.[11] An alternative approach employs semiparametric models based on nonparametric smooth functions to identify general nonlinear growth patterns.

[11] One difference is that Papageorgiou (2002) employs the threshold regression of Hansen (2000) while Tan (2010) employs a generalized regression tree algorithm.

Notable examples include Durlauf, Kourtellos, and Minkin (2001) and Mamuneas, Savvides, and Stengos (2006).

One difficulty with all the above studies is that they ignore the problem of endogeneity in the threshold variable. This is important because, as Kourtellos, Stengos, and Tan (2011) point out, if the threshold variable is endogenous, the above approaches will yield inconsistent parameter estimates for the regime-specific partial effects. A second difficulty is that the existing literature either deals with model uncertainty in the linear context or attempts to systematically uncover possible nonlinearity/heterogeneity, but approaches that coherently address both problems at the same have been lacking. Some initial attempts in this direction have been made by Brock and Durlauf (2001), Kourtellos, Tan, and Zhang (2007), and Cuaresma and Doppelhofer (2007).

A methodological contribution of this paper is to propose an econometric methodology that attempts to deal with both parameter heterogeneity and theory uncertainty by synthesizing model averaging and structural threshold regression, which has been recently proposed by Kourtellos, Stengos, and Tan (2011). A key innovation of the structural threshold regression (STR) model is that it allows for both the threshold variable as well as the (slope) regressors to be endogenous. In practice, taking into account the fact that the STR estimation is not affected by the model specification, we first consider the issue of theory uncertainty and thus, we estimate the 2SLS-BMA linear volatility growth model. Afterwards, using various model specifications derived directly from the 2SLS-BMA results, we test whether observations are organized into different volatility processes (i.e., volatility regimes) depending on whether the observed value for some threshold variable is above (or below) a critical/threshold value. Finally, we take into account again for theory uncertainty, by estimating the above models below and above the estimated threshold point using model averaging.

According to the results, we find strong evidence for a threshold under various model specifications and using different endogenous variables, especially for Initial Income and Public Debt. In the STR estimation, initial life expectancy and institutions are the two key variables which affect negatively growth volatility, especially in high Income, high Inflation Volatility, and high Debt countries. The negative effect of institutions on Growth Volatility is also identified for low Inflation Volatility countries and in countries with low Public Debt, the impact of Life Expectancy is also strong. Finally, in countries with greater Trade Openness the effect of Life Expectancy and Population Growth is negative and significant whereas the negative effect of institutions on Growth Volatility is identified in countries with trade barriers.

The paper is organized as follows. Section 4.2 describes our empirical methodology and Section 4.3 describes our data. In Section 4.4 we present the main results of the paper and finally, in Section 4.5, we conclude.

## 4.2 Empirical methodology

### 4.2.1 The canonical growth volatility model

Following the literature we define the volatility of economic growth, $\sigma_{g,i,t}$, as the standard deviation of the growth rate of real per capita GDP over the time interval $t-1$ to $t$. In this paper we present results for 10-year periods, which allow us to exploit the panel structure of the data as well as to estimate the standard deviation using time-series variation for each country.[12]

Then the canonical growth volatility regression takes the form:

$$\sigma_{g,it} = X'_{it}\beta + e_{it}, \qquad (4.2.1)$$

where $i = 1, 2, ..., N_t$ and $t = 1, 2, ..., T$. $X_{it}$ is a $k \times 1$ vector of variables that is partitioned into a $k_1 \times 1$ vector of exogenous/predetermined determinants, $X_{1it}$, and a $k_2 \times 1$ vector of right hand side endogenous determinants, $X_{2it}$. $e_{it}$ is an *i.i.d.* error term. Assuming a $l \times 1$ vector of instrumental variables $Z_{it} = (X'_{1it}, Z'_{2it})'$ such that $l \geq k$ the implied reduced form for $X_{2it}$ takes the following form

$$X_{2it} = \Gamma'_2 Z_{it} + V_{Xit}, \qquad (4.2.2)$$

where $V_{Xit}$ is a vector of *i.i.d.* errors. In this paper we instrument all endogenous variables using their lagged values. Equation (4.2.1) is then estimated by 2SLS.

[12] Different papers use different periods to estimate the standard deviation. For example, while Acemoglu, Johnson, Robinson, and Thaicharoend (2003) use a 27-year time interval to estimate the standard deviation Bekaert, Harvey, and Lundblad (2006) only use 5 years. For robustness we also used 20-year periods and find similar results.

Putting aside issues related to the endogeneity of growth volatility determinants and the validity of instruments we argue that the existing literature, which is based on equation (4.2.1) suffers from two important sources of model uncertainty - very much alike the problems that characterize the traditional growth regressions - theory uncertainty and parameter heterogeneity.

As discussed above, we propose an econometric methodology that unifies two recent econometric techniques; i.e., Bayesian Model Averaging and Structural Threshold Regression (STR), that will allow us to deal with the two problems of parameter heterogeneity and theory uncertainty simultaneously.

Next, we describe a STR model for growth volatility that deals with the problem of parameter heterogeneity alone. Then, we propose a model averaging approach for STR models to account for both theory uncertainty and parameter heterogeneity.

### 4.2.2 Threshold growth volatility model

We now describe the STR model by Kourtellos, Stengos, and Tan (2011), that allows for endogeneity in the slope regressors $X_{it}$ as well as the threshold variable.[13] This model can be viewed as a generalization of the simple threshold regression framework of Hansen (2000) and Caner and Hansen (2004) to allow for the endogeneity of the threshold variable and regime specific heteroskedasticity.

Consider a threshold variable $q_{it}$ such as public debt that can organize the observations into regimes and define the following indicator function

$$I(q_i \leq \gamma) = \begin{cases} 1 & \text{iff } q_i \leq \gamma : \text{Regime 1} \\ 0 & \text{iff } q_i > \gamma : \text{Regime 2} \end{cases} \tag{4.2.3}$$

and $I(q_i > \gamma) = 1 - I(q_i \leq \gamma)$. In this paper, we assume that $q_{it}$ can be any non-constant variable that belongs to the set of determinants $X_{it}$. We assume that $q_{it}$ is endogenous so

[13] The threshold model of Caner and Hansen (2004) (IVTR) allows only for the endogeneity of the slope regressors and maintains the assumption of the exogeneity of the threshold. STR reduces to IVTR when $\kappa = 0$.

that the reduced form equation that determines which regime applies takes the form

$$q_i = \pi_q' Z_i + v_{qi}, \tag{4.2.4}$$

It is worth noting that the above reduced form equation is analogous to the selection equation that appears in the literature on limited dependent variable models. However, there is one important difference. While in sample selection models, we observe the assignment of observations into regimes but the (threshold) variable that drives this assignment is taken to be latent, here, it is the opposite; we do not know which observations belong to which regime (i.e., we do not know the threshold value), but we can observe the threshold variable.

Following Kourtellos, Stengos, and Tan (2011) we can generalize (4.2.1) to allow for two regimes as follows:

$$\sigma_{g,it} = X_{it}'\beta + X_{it}' I(q_{it} \leq \gamma)\delta + \lambda_{it}(\gamma)\kappa + \varepsilon_{it}, \tag{4.2.5}$$

where $E(\varepsilon_{it}|Z_{it}) = 0$.

The term $\lambda_{it}(\gamma)$ is a scalar variable that involves an inverse Mills ratio term for each regime in order to restore the conditional mean zero property of the errors. In particular, $\lambda_{it}(\gamma)$ is defined as follows:

$$\lambda_{it}(\gamma) = \lambda_{1it}(\gamma) I(q_i \leq \gamma) + \lambda_{2it}(\gamma) I(q_{it} > \gamma),$$

with $\lambda_1(\gamma - Z_{it}'\pi_q) = -\frac{\phi(\gamma - Z_{it}'\pi_q)}{\Phi(\gamma - Z_{it}'\pi_q)}$ and $\lambda_2(\gamma - Z_{it}'\pi_q) = \frac{\phi(\gamma - Z_{it}'\pi_q)}{1-\Phi(\gamma - Z_{it}'\pi_q)}$. The functions $\phi(\cdot)$ and $\Phi(\cdot)$ are the normal pdf and cdf, respectively.

Finally, note that the coefficients $\beta$ are the coefficients of the second regime, that is $\beta = \beta_2$ and $\delta$ is the difference between the coefficients of regime 1, $\beta_1$ and regime 2, $\beta_2$; that is, $\delta = \beta_1 - \beta_2$. Equation (4.2.5) reduces to the linear growth volatility model in 4.2.1 when $\delta = \kappa = 0$.

The estimation of the threshold parameter is based on a concentrated least squares method while the slope estimates are obtained using 2SLS or GMM. The asymptotic distribution of the threshold parameter $\gamma$ is nonstandard as it involves two independent Brownian motions with two different scales and two different drifts. Confidence intervals are provided by an inverted likelihood ratio approach; see Kourtellos, Stengos, and Tan (2011).

### 4.2.3 Model averaging

Taking into account the fact that the STR estimation is not affected by the model specification, we first consider the issue of theory uncertainty explained in this section. In particular, we estimate the 2SLS-BMA linear volatility growth model, and afterwards, using various model specifications derived directly from the 2SLS-BMA results, we test whether observations are organized into different volatility processes (i.e., volatility regimes) depending on whether the observed value for some threshold variable is above (or below) a critical/threshold value. Finally, we take into account again for theory uncertainty, by estimating the above models below and above the estimated threshold point using model averaging.

Let us consider a set of potential threshold variables $q_s$, $s = 1, 2....Q$, which generates a set of regressors $\{X'_{it} I(q_{sit} \leq \gamma_s)\}_{s=1}^{Q}$ and a set of inverse of Mills ratio terms $\{\lambda_{sit}\}_{s=1}^{Q}$. To account for all possibilities we also need to account for the product $\{X'_{it} I(q_{sit} \leq \gamma_s)\}_{s=1}^{Q}$ over all possible combinations of threshold variables.[14] Then given threshold parameters $\gamma_s$, we can generalize equation (4.2.1) into a multiple threshold regression model

$$\begin{aligned} \sigma_{g,it} &= X'_{mit}\beta_m + \sum_{s=1}^{k_m^q} X'_{it} I(q_{sit} \leq \gamma_s)\delta_{ms} + \sum_{s=1}^{k_m^q} \lambda_{sit}(\gamma_s)\kappa_{ms} \\ &\quad + \sum_{s=1}^{k_m^q} X'_{it} \prod_{j \in Q} I(q_{jit} \leq \gamma_j)\zeta_{ms} + \varepsilon_{mit}, \end{aligned} \tag{4.2.6}$$

where the subscript $m$ denotes a particular combination of regressors in equation (4.2.6) and corresponding slope coefficients $\theta_m = (\beta'_m, \delta^{*'}_m, \zeta^*_m, k^*_m)'$, where $\delta^*_m = (\delta'_1, ...., \delta'_{Q_m})'$, $\zeta^*_m = (\zeta'_1, ...., \zeta'_{Q_m})'$ and $k^*_m = (\kappa_1, ...., \kappa_{Q_m})'$. The parameter $\zeta^*_m$ captures the effect of the simultaneous presence of $Q_m$ thresholds effects and the product, $\prod_{j \in Q}$, is taken over all possible combinations of the $Q$ threshold variables. A sufficient identification condition is that the data matrix of the instrumental variables of the largest model has full rank. Finally, we can estimate $\theta_m$ conditional on the available data, $D$, and the specification of the growth volatility model, $M_m$, by 2SLS or GMM, $\widehat{\theta}_m^{2SLS}$.

[14]For example in the case of $X_{it} = 1$ and two threshold variables $q_1$ and $q_2$, the regressors are given by 1, $I(q_1 \leq \gamma_1)$, $I(q_2 \leq \gamma_2)$, and $I(q_1 \leq \gamma_1) \cdot I(q_2 \leq \gamma_2)$.

The set of all possible combinations of regressors from this set constitutes the model space, denoted by $M$. For simplicity we only consider just-identified systems. This implies that for any given $M_m$ we can obtain an associated first stage model given by model specific versions of equations (4.2.2) and (4.2.4). Given that the true model is in the model space, $M$, we can think of each model $M_m$ as a model that places linear restrictions on the largest model. An important finding in Kourtellos, Stengos, and Tan (2011) (see, Remark 1 of Proposition 4.1) is that the estimate of the threshold value from the restricted model, and the threshold value estimate from the unrestricted model, both converge to the true threshold value, asymptotically. The finding that the threshold estimate for the restricted model is a consistent estimator for gamma is therefore particularly useful when we do not know what the true model is due to theory uncertainty. Kourtellos, Stengos, and Tan (2011) also show that the estimator of the threshold parameter is super-consistent while the slope estimators, are root-n consistent and hence the slope parameters, $\theta_m$, can be estimated as if the threshold parameters were known.

A standard approach to estimate equation 4.2.6, is to estimate a baseline model which is closest to the theory and then report a few robustness exercises that include some additional controls. An alternative method to evaluate the relative evidentiary support of competing theories includes a large number of variables and those variables that prove to be significant are then rendered as the important determinants. This approach is often referred as a kitchen sink approach. However, both approaches do not systematically address the problem of model uncertainty and do not provide robust evidence but rather rely on strong priors of the econometrician. As Brock and Durlauf (2001) and others have argued, the inherent openendedness of new growth theories presents unique challenges to researchers in exploring their quantitative consequences on growth. Because the inclusion of one set of growth theories says nothing about whether other possible growth theories should be included (or not) in the model, growth researchers face substantial model uncertainty in their work. The fear is that the inclusion or exclusion of variables may significantly alter the conclusions one had previously arrived at for, say, the effect of institutions on Growth Volatility based on a particular model in the model space. In this case, the model space refers to the set of all possible models generated by the set of growth regressors.

Then, how can we obtain robust determinants in equation (4.2.6) and more generally robust inference about the structural parameters $\theta_m$ that do not condition on the model choice? We do so by employing a Bayesian Model Averaging (BMA) approach by constructing estimates conditional not on a single model, but on a model space whose elements span an appropriate range of determinants suggested by a large body of work. A number of recent papers have documented the advantages of using BMA in constructing robust estimates primarily in the context of the linear model. See for example, Brock and Durlauf (2001), Fernandez, Ley, and Steel (2001), Sala-I-Martin, X. and Doppelhofer, G. and Miller, R. (2004), and Durlauf, Kourtellos, and Tan (2008), Masanjala and Papageorgiou (2008).

In particular, we employ the 2SLS-BMA approach proposed by Durlauf, Kourtellos, and Tan (2011) and Eicher, Lenkoski, and Raftery (2010) that computes the weighted average of model-specific estimates using 2SLS estimates; the weights are constructed to be analogous to posterior model probabilities. The model average estimator for the slope parameters takes the following form

$$\widehat{\theta}_{M}^{2SLS} = \sum_{m \in M} \widehat{\mu}(M_k|D)\widehat{\theta}_{m}^{2SLS} \tag{4.2.7}$$

For the model weights, $\widehat{\mu}(M_k|D)$ we use Bayes' rule, so that each weight is the product of the integrated likelihood of the data given a model, $\widehat{\mu}(D|M_k)$, and the prior probability for a model, $\mu(M_k)$:

$$\widehat{\mu}(M_k|D) \propto \widehat{\mu}(\theta|M_k)\mu(M_k) \tag{4.2.8}$$

As standard in the literature, we assume a uniform model prior so that the prior probability that any variable is included in the true model is taken to be 0.5. The integrated likelihood of model $M_k$ reflects the relative goodness of fit of different models and is approximated by the Bayesian information criterion (BIC).

We also compute the corresponding standard errors using the posterior variance of $\theta$

$$\widehat{V}_{D,M}^{\theta} = \sum_{k \in M} \widehat{V}_{D,k}^{\theta}\widehat{\mu}(M_k|D) + \sum_{k \in M}(\widehat{\theta}_{D,k}^{2SLS} - \widehat{\theta}_{D,M^S}^{2SLS})^2\widehat{\mu}(M_k|D) \tag{4.2.9}$$

where $\widehat{V}_{D,k}^{\theta}$ is the model-specific posterior variance of the partial likelihood estimator estimator. The first term in equation (4.2.9) is the average of the posterior variances within models and the second term is the variance of the posterior means across models (i.e. weighted average of the squared deviations of the model-specific from the model averaged estimates). We also report the posterior probability of inclusion for each covariate, which is the sum of the posterior probability of all the models for which that variable appears. It is meant to capture the (posterior) probability that that covariate is in the true model after looking at the data.[15]

## 4.3 Data

We employ an unbalanced 10-year period panel dataset covering 28 countries in 1970-79, 87 in 1980-89, 95 in 1990-99, and 123 in 2000-2009 with a total of 333 observations. The dependent variable is computed as the standard deviation of the growth rate of real per capita GDP over the time interval sampled from PWT 7.0.

We next describe the variables that generate the model space, $M$. We should also note that in all exercises we include a constant and a time trend. In the absence of strong theoretical guidance we follow the existing empirical literature and we take the stand that the processes of growth volatility and economic growth share the same information set. Therefore, the set of possible theories and their proxies that have been proposed in the empirical growth literature can also be used in the context of growth volatility.

We start with the Solow variables, which include the logarithm of population growth plus 0.05 (*Population Growth*), the logarithm of the average investment to GDP ratio (*Investments*), the logarithm of the average years of secondary and tertiary schooling for male population over 25 years of age (*Schooling*), and the logarithm of real GDP per worker (*Initial Income*). Theoretical work by Acemoglu and Zilibotti (1997) and Koren and Tenreyro (2007) suggest a negative relationship between growth volatility and initial income.

As argued in Acemoglu, Johnson, Robinson, and Thaicharoend (2003) the traditional macroeconomic argument links growth volatility to bad macroeconomic policies. To account for the effect of macroeconomic policy on volatility we use the logarithm of average inflation rate (*Inflation Rate*), the standard deviation of the Inflation Rate (*Inflation Volatility*), the ratio of exports plus imports to GDP (*Openness*)), the ratio of government consumption to GDP (*Government*), and the logarithm of the average public debt to GDP (*Debt*).

Following the recent literature in economic growth that emphasizes the role of



[15] As a final note this approach can be viewed as a "hybrid" approach to model averaging in the sense that we mix frequentist probability statements about observables given unobservables and Bayesian probability statement about unobservables given observables; see Durlauf, Kourtellos, and Tan (2011) for more details.

fundamental determinants we include measures of institutions, geography, and health. As argued by Acemoglu, Johnson, Robinson, and Thaicharoend (2003) countries with weak institutions are more likely to experience high volatility. Hence, we include constraints placed on the executive (*Executive Constraints*), which measures institutional and other constraints that are placed on presidents and dictators (or monarchies). For geography, Malik and Temple (2009) found evidence that geographical characteristics of countries have effects on growth volatility. A key finding by Malik and Temple is that remoteness is associated with a lack of export diversification, which generates high volatility in the terms of trade and in output. Therefore we include a climate variable as well as a variable that measures geographic isolation. The climate variable we use is the percentage of a country's land area classified as tropical and subtropical via the Koeppen-Geiger system (*Tropics*) while the geographic isolation proxy is the percentage of a country's land area within 100km of an ice-free coast (*LCR100KM*). For health we include the logarithm of the average of life expectancy (*Life Expectancy*), which was found to be an important determinant of growth volatility by Bekaert, Harvey, and Lundblad (2006).

Finally, we include variables that capture regional heterogeneity, which consists of dummy variables for *East Asia*, *Sub-Saharan African*, and *Latin American and Caribbean* countries.

Table 4.1 presents summary statistics for the pooled data. The variables are drawn from various sources. A detailed description of the variables and their sources is given in Table B1 of the Appendix.

## 4.4 Results

### 4.4.1 Results for Linear Volatility Model

We start our analysis by investigating the linear volatility growth model described by equation (4.2.1). Table 4.2 presents the 2SLS-BMA results while Table 4.3 shows the LS-BMA results. The first three columns of each Table present the model averaging results. The first column shows the posterior probability that each of the covariates is included in the true Growth Volatility model, while the second and third columns present the BMA posterior means and standard errors for each covariate, respectively. The model averaging results are compared with two individual models of the model space, which are typically reported in empirical exercises. These are the posterior mode model; i.e., the best model in terms of posterior weight, and the largest model; i.e., the model with all the covariates that span the model space. The largest model is potentially of interest because it is a low-bias model but with possibly high variance. In Table 4.2 we posit that there is no uncertainty in the first stage, and assuming just identification, we instrument all the endogenous variables with their lag values, which are predetermined. This is standard in the literature and note that in all cases the F-statistics from the first stage, are well above 10 that suggest that our instruments are not weak.

The 2SLS-BMA analysis yields the following results: First, we find that Population Growth, Life Expectancy, and Executive Constraints are the most robust determinants of Growth Volatility with very high posterior inclusion probabilities at 87.4%, 90.1% and 94%, respectively, well above the 50% prior inclusion probability. Second, in terms of significance, the corresponding posterior means for Population Growth and Life Expectancy are negative and not significant, but the posterior mean for Executive Constraints is negative and significant at the 5% level. Finally, with the exception of Trade Openness, the Sub-Saharan Africa and the Latin America dummies (with posterior inclusion probabilities 59.6%, 56.5% and 53.9% respectively, and not significant), all the other variables have posterior inclusion probability well below the 50% prior.



Our findings confirm the results of Bekaert, Harvey, and Lundblad (2006) and Acemoglu, Johnson, Robinson, and Thaicharoend (2003). As Acemoglu, Johnson, Robinson, and Thaicharoend (2003) point out, the major causes of the large cross-country differences in volatility are institutional, and none of the standard macroeconomic policy variables appear to be the primary mediating channels through which institutional causes lead to economic instability. Instead, these macroeconomic problems are symptoms of deeper institutional causes. Furthermore, low probability of survival reduces any incentives for capital accumulation and affects growth and growth volatility via lower levels of physical and human capital investments, a result which was also confirmed by Bekaert, Harvey, and Lundblad (2006).

Many researchers have also highlighted the role of these determinants in the economic growth context. In particular, Durlauf, Kourtellos and Tan (2008, 2011) using again BMA analysis and a similar mode space, they show that Initial Income, Government Consumption, Inflation, demography and geography, have a negative effect on economic growth, whereas the effect of Investments, Openness, and institutions is positive.

The BMA results are confirmed by the results from the posterior mode model: Population Growth, Life Expectancy and Executive Constraints are included in the model, with no significance for Life Expectancy and significance at the 5% and 1% for Population Growth and institutions, respectively. In this model, the Sub-Saharan Africa and the Latin America dummies are also included, both positive and significant at the 1% level.

As in the posterior mode model, we find that GDP per capita growth is more volatile in Sub-Saharan and in Latin American countries compared to those in the rest of the world, when we consider the largest model. Furthermore, Population Growth and institutions affect negatively Growth Volatility (at the 5% level), in contrast to Trade Openness which increases Growth Volatility (at the 5% level).

The above findings are also confirmed in the BMA-LS analysis presented in Table 4.3. In particular, as in the BMA-2SLS case, only Population Growth, Life Expectancy and Executive Constraints have high posterior inclusion probability (91.5%, 97.5% and 95.3%, respectively), whereas in terms of significance, all variables are negative and significant at the 10% for Population Growth, and at the 5% for Life Expectancy and Executive Constraints. The remaining variables - except from Trade Openness- have a posterior inclusion probability below the 50% prior.

The posterior mode model from the BM-LS analysis confirms the earlier findings: Population Growth is negative and significant determinant at the 5% level, whereas Trade Openness, Life Expectancy and Executive Constraints are also included with significance at the 1% level. In particular, higher levels of Trade Openness, along with bad institutions and lower Life Expectancy, increase Growth Volatility. The results from the largest model in this section, are compatible with the largest model from the BMA-2SLS analysis.

### 4.4.2 Results for STR Volatility Model

We next consider whether observations are organized into different volatility processes (i.e., volatility regimes) depending on whether the observed value for some threshold variable is above (or below) a critical/threshold value. That is, we model potential heterogeneity in the growth volatility process using the STR model by Kourtellos, Stengos, and Tan (2011). A key innovation of the STR model is that it allows endogenous variables to be employed as threshold variables. Our set of candidate threshold variables include Initial Income, Population Growth, Investments, Schooling, Debt, Government, Inflation, (log) Inflation

Volatility,(log) Trade Openness, (log) Life Expectancy and Executive Constraints.

In Table 4.4 we test the null hypothesis of a linear model against the alternative of a threshold for each candidate threshold variable, $H_0 : \delta_s = 0,\ s = 1, 2....Q$. We do so by employing the sup Wald test of Kourtellos, Stengos, and Tan (2011), which is an extension of the Davies (1977) Sup test to the GMM framework.[16] Since the threshold parameter, $\gamma_s$, is not identified under the null hypothesis of a linear model (i.e. no threshold effect), the p-values are computed by a bootstrap method, which relies on the arguments of Hansen (1996). Specifically, the p-values are computed by a bootstrap that fixes the regressors from the right-hand side of equation (4.2.5) and generating the bootstrap dependent variable from the distribution $N(0, \widehat{\varepsilon}_{it}^2)$, where $\widehat{\varepsilon}_{it}$ is the demeaned residual from the estimated STR model.

In terms of the set of slope regressors $X$, we consider four possibilities: (i) includes the set of regressors with posterior inclusion probability greater than 75% in the linear 2SLS-BMA exercise, following the general rule developed by Jeffreys (1961) and refined by Kass and Raftery (1995). In particular, according to this rule, probabilities smaller than 50% are seen as evidence against an effect, and the evidence for an effect is either weak, positive, strong, or decisive for posterior probabilities ranging from 50-75%, 75-95%, 95-99%, and greater than 99%, respectively. This set thus, includes Population Growth, Life Expectancy and Executive Constraints. We take this case to be our baseline model (see Table 4.4, Model 1). For robustness we also consider the case (ii) drops the variables that are not statistically significant in (i) i.e., we are left with Executive Constraints (see Table 4.4, Model 3). We also consider (iii) considers the set of covariates in the (linear) posterior mode model; i.e., the variables in (i) plus the Sub-Saharan Africa and the Latin America dummies (see Table 4.4, Model 2). Finally, (iv) we include the full set of regressors; i.e., the largest model (see Table 4.4, Model 4). In all cases, we include a constant and trend.

Using 1000 bootstrap replications, we present in Table 4.4 the SupW statistic along with the corresponding bootstrap p-value for each candidate threshold variable for each of the cases (i)-(iv). The p-values for case (i) (Model 1) show that the null hypothesis of a linear model is rejected at the 1% level for our baseline model for Initial Income, Population Growth, Public Debt, Inflation Volatility and Trade Openness. The results are also supported if we consider cases (ii)-(iv). Overall, under any model specification the results support the

[16]For robustness purposes we also employed the threshold sup tests by Caner and Hansen (2004) and Hansen (2000) that ignore the issue of the endogeneity in the threshold variable and generally found similar results.

existence of a threshold, especially for initial income and public debt.

In the next section, we investigate the influence of every determinant on growth volatility by splitting the sample based on the estimated threshold points discussed earlier. Table 4.5 presents the estimation results for the STR growth volatility model in equation (2.5) for the baseline model (case (i)). Recall that for this specific model there is evidence for a threshold for Initial Income, Population Growth, Public Debt, Inflation Volatility and Trade Openness. We report the estimated threshold point and the 90% confidence interval for each of these threshold variables. Each of these threshold variables would therefore split the full set of observations into two subsamples (regimes). For each regime, we present the associated coefficient estimates and standard errors for the regressors from the baseline model. Table 4.6 then shows the country breakdowns for each of the two regimes corresponding to the respective threshold variables.

According to the results, in countries with high Initial Income (above 6.720), higher Initial Life Expectancy and better institutions have a strong (1%) negative effect on Growth Volatility. Also, higher Population Growth affects also negatively (at 10%) Growth Volatility for countries with high initial income. These results add to the previous findings in the sense that they confirm that Population Growth, Life Expectancy and Institutions are strong determinants of Growth Volatility, especially for high Income countries.

In addition, in countries with high Public Debt (above 4.2958) Initial Life Expectancy (10%) and Executive Constraints (5%) affect negatively Growth Volatility. However, in countries with low Public Debt, the impact of Life Expectancy is more profound (1%). In countries with high Inflation Volatility (above 1.3696) again, Life Expectancy (5%) and institutions (10%) affect negatively Growth Volatility. The negative effect of institutions on Growth Volatility is also identified for low Inflation Volatility countries (at 10%). Finally, in countries with greater Trade Openness (above 4.3975) the effect of Life Expectancy and Population Growth is negative and significant (at 1% and 10% respectively), whereas the negative effect of institutions (at 1%) on Growth Volatility is identified in countries with trade barriers.

In order to account for model uncertainty, we also estimate the above model using model averaging. To do so, we exploit the framework described in equation 4.2.6. Table 4.7 presents the model averaging results using LS-BMA. The first column of Table 4.7 shows the posterior probability that each of the covariates is included in the true Growth Volatility model, while the second and third columns present the BMA posterior means and standard errors for each covariate. The remaining columns show the estimation results for the posterior mode and the largest model, respectively. The model space includes the covariates of our basic model (Population Growth, Executive Constraints, Life Expectancy), but also, the Inverse Mills Ratios for each regime derived directly from the STR analysis in order to restore the conditional mean zero property of the errors,and of course, an indicator function for every threshold which represents the regimes.

According to the results, we find that institutions (98.7%) affect Growth Volatility negatively and significantly at the 1% level, in countries with high Debt, high Population Growth, high Inflation Volatility, high Trade Openness and high Initial Income. For countries however characterized by high values for all threshold variables except from Debt, the effect of Executive Constraints (100%) is positive and significant at the 1% level.

Population Growth in low Debt (99.2 %) countries but with high Population Growth, Inflation Volatility Trade Openness and Initial Income, has positive and significant effect on Growth Volatility (1%). Nevertheless, Population Growth affects negatively Growth Volatility (98.9% and 1% significance) in countries with low Population Growth, but with high values in the remaining threshold variables. Important also is the effect of Population Growth in countries with low Inflation volatility (but high Debt, Population Growth, Openness and Initial Income) and low Initial Income (but high Debt, Population Growth, Inflation Volatility and Trade Openness). Despite the fact that both variables have high posterior inclusion probabilities (85.1 % and 88.2% , respectively) they are not statistically significant. Finally, Life Expectancy has a negative effect on Growth Volatility in countries with high values for all threshold variables except from Population Growth (96.4% and at 1% level significant). The earlier findings are also confirmed by the posterior mode model.

In Table 4.8 we extend the model by including the full set of thresholds, ignoring the results from the bootstrap SupW statistic from the STR analysis. According to these results, again, Executive Constraints affects positively Growth Volatility in low Debt countries but with high values in the remaining threshold variables (with 100% PIP and at the 1% level).

In countries with low Investments but high values in the remaining threshold variables, Life Expectancy (98.8%) and Population Growth (97.9%) have a negative effect on Growth Volatility at the 1% level.

## 4.5 Conclusion

In this paper we uncover growth volatility regimes and identify their robust determinants using a large international panel of countries over the period 1970-2009. In order to identify any regimes and deal with model uncertainty, we employ two recent econometric techniques: Bayesian Model Averaging and Structural Threshold Regression (STR). According to the results, the null hypothesis of a linear model is rejected strongly for our basic model for Initial Income, Population Growth, Public Debt, Inflation Volatility and Trade Openness. Overall, under any model specification the results support the existence of a threshold, especially for initial income and public debt. In the STR estimation, countries with high Initial Income higher Initial Life Expectancy and better institutions have a strong negative effect on Growth Volatility. Higher Population Growth affects also negatively Growth Volatility for countries with high Initial Income. In addition, in countries with high Public Debt, Initial Life Expectancy and Executive Constraints affect negatively Growth Volatility. However, in countries with low Public Debt, the impact of Life Expectancy is more profound. In countries with high Inflation Volatility again, Life Expectancy and institutions affect negatively Growth Volatility. The negative effect of institutions on Growth Volatility is also identified for low Inflation Volatility countries. Finally, in countries with greater Trade Openness the effect of Life Expectancy and Population Growth is negative and significant whereas the negative effect of institutions on Growth Volatility is identified in countries with trade barriers.

Our findings can help for better understanding of the development process, as well as, for the design of economic policies aiming at stabilizing the growth path especially in underdeveloped economies.

## Table 4.1: Descriptive Statistics

This table presents the summary statistics for the pooled sample over the period 1970-2009.

| | Mean | St. Dev. | Min | Max |
|---|---|---|---|---|
| Growth Volatility | 0.0430 | 0.0320 | 0.0059 | 0.3126 |
| Initial Income | 8.3710 | 1.2484 | 5.8620 | 10.711 |
| Lag of Initial Income | 8.2712 | 1.2266 | 5.5774 | 10.560 |
| Population Growth | -2.7247 | 0.1745 | -3.229 | -2.365 |
| Lag Population Growth | -2.7007 | 0.1881 | -3.243 | -1.995 |
| Investments | 3.0463 | 0.3621 | 1.8732 | 3.8915 |
| Lag Investments | 3.0243 | 0.4068 | 1.5317 | 4.3127 |
| Schooling | 0.5583 | 0.8352 | -2.3502 | 2.1252 |
| Lag Schooling | 0.2860 | 0.9505 | -2.9136 | 1.9010 |
| Debt | 3.9479 | 0.6768 | 1.6245 | 6.3274 |
| Lag Debt | 3.8095 | 0.8050 | 1.1170 | 6.4624 |
| Government | 2.2128 | 0.4540 | 1.0562 | 3.5609 |
| Lag Government | 2.2121 | 0.4953 | 0.9211 | 3.6945 |
| Inflation | 2.2757 | 1.0650 | -1.9518 | 7.5714 |
| Lag Inflation | 2.3024 | 1.2059 | -1.4595 | 8.2583 |
| Inflation Volatility | 65.683 | 432.29 | 0.3276 | 4429.1 |
| Lag Inflation Volatility | 55.254 | 454.57 | 0.3450 | 5774.8 |
| Openness | 67.918 | 37.459 | 9.7683 | 199.86 |
| Lag Openness | 61.689 | 37.405 | 8.4039 | 265.57 |
| Life Expectancy | 64.124 | 10.554 | 32.827 | 81.076 |
| Lag Life Expectancy | 62.643 | 10.566 | 30.473 | 79.536 |
| Executive Constraints | 4.8433 | 2.0765 | 1.0000 | 7.0000 |
| Lag Executive Constraints | 4.4704 | 2.2747 | 1.0000 | 7.0000 |
| Tropics | 0.4233 | 0.4294 | 0.0000 | 1.0000 |
| LCR100KM | 0.4634 | 0.3631 | 0.0000 | 1.0000 |

### Table 4.2: Linear growth volatility model using 2SLS-BMA

The table presents 2SLS-BMA results for the linear growth volatility model in equation (4.2.1). The instruments are lagged values of the variables. The posterior inclusion probability (PIP) of a variable is the sum of the posterior probabilities models that include that variable. The posterior mean (PM) is the average of the coefficient estimates (COEF) of individual models weighted by posterior probability. The posterior standard error (PSE) is the BMA estimate for the standard error (SE) taking model uncertainty into account. All reported standard errors refer to heteroskedasticity consistent standard errors.

| | Model averaging | | | Posterior mode | | Largest | |
|---|---|---|---|---|---|---|---|
| | PIP | PM | PSE | COEF | SE | COEF | SE |
| Constant | 1.0000 | 0.0180 | 0.0593 | -0.0263 | 0.0385 | -0.0233 | 0.0484 |
| Trend | 1.0000 | -0.0010 | 0.0019 | -0.0007 | 0.0015 | -0.0020 | 0.0020 |
| EASIA | 0.0499 | 0.0001 | 0.0017 | - | - | 0.0059 | 0.0064 |
| LAC | 0.5396 | 0.0060 | 0.0063 | 0.0120 | 0.0034 | 0.0134 | 0.0056 |
| SSA | 0.5656 | 0.0116 | 0.0114 | 0.0212 | 0.0060 | 0.0226 | 0.0063 |
| Initial Income | 0.0214 | 0.0000 | 0.0005 | - | - | -0.0006 | 0.0034 |
| Population Growth | 0.8742 | -0.0275 | 0.0193 | -0.0406 | 0.0169 | -0.0337 | 0.0156 |
| Investments | 0.0616 | 0.0002 | 0.0019 | - | - | 0.0025 | 0.0063 |
| Schooling | 0.1111 | 0.0004 | 0.0020 | - | - | 0.0054 | 0.0050 |
| Debt | 0.0172 | 0.0000 | 0.0004 | - | - | -0.0015 | 0.0039 |
| Government | 0.0258 | 0.0002 | 0.0014 | - | - | 0.0082 | 0.0054 |
| Inflation | 0.1253 | 0.0005 | 0.0016 | - | - | 0.0041 | 0.0034 |
| Inflation Volatility | 0.0252 | 0.0000 | 0.0000 | - | - | 0.0000 | 0.0000 |
| Openness | 0.5968 | 0.0001 | 0.0001 | - | - | 0.0001 | 0.0001 |
| Life Expectancy | 0.9016 | -0.0007 | 0.0005 | -0.0004 | 0.0003 | -0.0007 | 0.0007 |
| Executive Contraints | 0.9407 | -0.0037 | 0.0018 | -0.0044 | 0.0015 | -0.0038 | 0.0017 |
| Tropics | 0.0303 | 0.0000 | 0.0008 | - | - | -0.0045 | 0.0053 |
| LCR100KM | 0.0188 | -0.0001 | 0.0009 | - | - | 0.0007 | 0.0065 |

**Table 4.3: Linear growth volatility model using LS-BMA**

The table presents LS-BMA results for the linear growth volatility model in equation (4.2.1). The posterior inclusion probability (PIP) of a variable is the sum of the posterior probabilities models that include that variable. The posterior mean (PM) is the average of the coefficient estimates (COEF) of individual models weighted by posterior probability. The posterior standard error (PSE) is the BMA estimate for the standard error (SE) taking model uncertainty into account. All reported standard errors refer to heteroskedasticity consistent standard errors.

| | Model averaging | | | Posterior mode | | Largest | |
|---|---|---|---|---|---|---|---|
| | PIP | PM | PSE | COEF | SE | COEF | SE |
| Constant | 1.0000 | 0.0169 | 0.0483 | 0.0339 | 0.0277 | -0.0324 | 0.0376 |
| Trend | 1.0000 | -0.0006 | 0.0019 | -0.0007 | 0.0017 | -0.0012 | 0.0024 |
| EASIA | 0.0455 | 0.0001 | 0.0014 | - | - | 0.0083 | 0.0049 |
| LAC | 0.4196 | 0.0045 | 0.0059 | - | - | 0.0152 | 0.0042 |
| SSA | 0.4553 | 0.0069 | 0.0089 | - | - | 0.0190 | 0.0069 |
| Initial Income | 0.0475 | 0.0001 | 0.0011 | - | - | 0.0023 | 0.0041 |
| Population Growth | 0.9158 | -0.0367 | 0.0192 | -0.0364 | 0.0159 | -0.0422 | 0.0141 |
| Investments | 0.0677 | 0.0004 | 0.0020 | - | - | 0.0054 | 0.0048 |
| Schooling | 0.1168 | 0.0006 | 0.0022 | - | - | 0.0055 | 0.0041 |
| Debt | 0.0311 | 0.0001 | 0.0005 | - | - | 0.0019 | 0.0025 |
| Government | 0.0462 | 0.0002 | 0.0011 | - | - | 0.0032 | 0.0044 |
| Inflation | 0.1919 | 0.0006 | 0.0015 | - | - | 0.0020 | 0.0022 |
| Inflation Volatility | 0.0642 | 0.0000 | 0.0000 | - | - | 0.0000 | 0.0000 |
| Openness | 0.7612 | 0.0001 | 0.0001 | 0.0001 | 0.0000 | 0.0001 | 0.0001 |
| Life Expectancy | 0.9751 | -0.0011 | 0.0005 | -0.0013 | 0.0003 | -0.0014 | 0.0007 |
| Executive Contraints | 0.9536 | -0.0034 | 0.0014 | -0.0033 | 0.0011 | -0.0040 | 0.0013 |
| Tropics | 0.0341 | -0.0001 | 0.0009 | - | - | -0.0055 | 0.0047 |
| LCR100KM | 0.0242 | 0.0000 | 0.0010 | - | - | 0.0025 | 0.0052 |

## Table 4.4: Threshold tests

This table presents supWald tests for the null of hypothesis the linear model growth volatility model in equation (4.2.1) against the alternative hypothesis of the threshold model in equation (4.2.5). We consider four different models for the slope variables $X$ based on the 2SLS-BMA results of Table 4.2. Model 1 includes those variables that achieve at least PPI$\geq$ 75%, i.e. Population Growth, Life Expectancy, and Executive Constraints; Model 2 includes the variables of the posterior model, i.e. the variables of Model 1 plus the Sub-Saharan Africa and the Latin America dummies; Model 3 includes only statistical significant variable of Model 1, i.e. Executive Constraints; and Model 4 corresponds to the largest model and includes all the variables. All models include constant and trend.

| | Model 1 | | Model 2 | | Model 3 | | Model 4 | |
|---|---|---|---|---|---|---|---|---|
| | statistic | p-value | statistic | p-value | statistic | p-value | statistic | p-value |
| Threshold variable | | | | | | | | |
| Initial Income | 31.7152 | 0.0020 | 43.1362 | 0.0500 | 31.5736 | 0.0280 | 70.6119 | 0.0340 |
| Population Growth | 22.7686 | 0.0000 | 19.0444 | 0.1330 | 28.6539 | 0.0060 | 26.0636 | 0.3510 |
| Investments | 8.7673 | 0.6190 | 27.6877 | 0.0240 | 18.2073 | 0.0520 | 34.6254 | 0.4920 |
| Schooling | 14.1301 | 0.2240 | 15.1285 | 0.5850 | 17.7559 | 0.0580 | 26.8452 | 0.3810 |
| Debt | 23.6673 | 0.0040 | 34.7814 | 0.0090 | 24.8225 | 0.0030 | 41.3602 | 0.0510 |
| Government | 11.6638 | 0.3440 | 19.8855 | 0.1970 | 11.8249 | 0.3910 | 26.7643 | 0.2000 |
| Inflation | 12.1088 | 0.2650 | 12.6355 | 0.8000 | 16.1973 | 0.2390 | 208.8370 | 0.0190 |
| Inflation Volatility (log) | 39.5210 | 0.0000 | 29.7792 | 0.2450 | 42.8252 | 0.0390 | 54.4759 | 0.3820 |
| Openness (log) | 35.5853 | 0.0000 | 25.0009 | 0.0670 | 28.5092 | 0.0010 | 34.8930 | 0.3730 |
| Life Expectancy (log) | 13.7552 | 0.2640 | 17.1110 | 0.4280 | 17.6186 | 0.3590 | 52.2920 | 0.2230 |
| Executive Constraints | 11.3689 | 0.3180 | 72.6192 | 0.0030 | 11.7169 | 0.2190 | 3488.0000 | 0.0010 |

## Table 4.5: STR estimation of the baseline model

This table presents the estimation of the STR growth volatility model in equation (4.2.5) for the baseline model, which includes Population Growth, Life Expectancy, Executive Constraints, constant and trend. Here, we only consider the threshold variables for which the null hypothesis of the linear model is rejected. For the threshold variables of Openness and Inflation Volatility we use their logarithmic transformation. The (L) refers to GMM estimates of the linear growth volatility, which includes observations that correspond to the threshold variable with values smaller than the threshold estimate. Similarly, the (H) refers to GMM estimates of the linear growth volatility, which includes observations that correspond to the threshold variable with values larger than the threshold estimate. The JSSE refers to joint sum of squared errors and JSTAT refers to the J-statistic of the STR estimation.

| Threshold variable | **Initial Inc.** | | **Pop. Growth** | | **Debt** | | **Inf. Vol.** | | **Openness** | |
|---|---|---|---|---|---|---|---|---|---|---|
| Threshold Estimate | 6.7200 | | -2.7537 | | 4.2958 | | 1.3696 | | 4.3975 | |
| 90% CI | [6.71 6.731] | | [-2.7557 -2.6717] | | [3.5328 4.7208] | | [1.0706 2.1266] | | [4.2365 4.4675] | |
| Regimes | **(L)** | **(H)** | **(L)** | **(H)** | **(L)** | **(H)** | **(L)** | **(H)** | **(L)** | **(H)** |
| Constant | -0.3692 | 0.0737 | -0.7872 | -0.4365 | 0.0680 | 0.0321 | 0.0036 | 0.0521 | 0.1149 | 0.1110 |
| | (0.4803) | (0.0429) | (1.2588) | (1.1833) | (0.0390) | (0.1028) | (0.0415) | (0.0541) | (0.0419) | (0.0776) |
| Trend | 0.0133 | 0.0015 | -0.0007 | 0.0009 | 0.0013 | -0.0023 | 0.0049 | 0.0024 | 0.0023 | -0.0074 |
| | (0.0201) | (0.0015) | (0.0030) | (0.0029) | (0.0015) | (0.0051) | (0.0016) | (0.0021) | (0.0019) | (0.0036) |
| Population Growth | -0.1254 | -0.0307 | -0.2721 | -0.2771 | -0.0132 | -0.0413 | -0.0063 | -0.0256 | 0.0161 | -0.0408 |
| | (0.1968) | (0.0165) | (0.3032) | (0.5882) | (0.0169) | (0.0408) | (0.0187) | (0.0222) | (0.0187) | (0.0216) |
| Life Expectancy | 0.0023 | -0.0015 | -0.0014 | -0.0008 | -0.0008 | -0.0009 | 0.0000 | -0.0009 | -0.0001 | -0.0024 |
| | (0.0017) | (0.0005) | (0.0023) | (0.0010) | (0.0002) | (0.0005) | (0.0003) | (0.0004) | (0.0003) | (0.0009) |
| Executive Constraints | -0.0186 | -0.0036 | -0.0054 | -0.0059 | -0.0012 | -0.0090 | -0.0030 | -0.0032 | -0.0062 | 0.0015 |
| | (0.0173) | (0.0014) | (0.0057) | (0.0051) | (0.0014) | (0.0038) | (0.0017) | (0.0019) | (0.0016) | (0.0029) |
| Inverse Mills Ratio term | -0.0229 | | -0.2362 | | 0.0176 | | -0.0099 | | -0.0016 | |
| | (0.0125) | | (0.3869) | | (0.0109) | | (0.0039) | | (0.0106) | |
| JSSE | 0.2686 | | 0.2514 | | 0.2659 | | 0.2668 | | 0.2513 | |
| JSTAT | 0.0000 | | 0.0000 | | 0.0000 | | 0.0000 | | 0.0000 | |
| Observations | 33 | 300 | 129 | 204 | 235 | 98 | 134 | 199 | 226 | 107 |

# Table 4.6: Regimes for various threshold variables

This table presents the countries sorted in regimes according to various thresholds. (H) denotes the High regime while (L) denotes the Low regime.

| Threshold | *Initial Income* | | | | *Pop. Growth* | | | | *Debt* | | | | *Openness* | | | | *Inf. Vol.* | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (1) | (2) | (3) | (4) | (1) | (2) | (3) | (4) | (1) | (2) | (3) | (4) | (1) | (2) | (3) | (4) |
| **Europe** | | | | | | | | | | | | | | | | | | | | |
| Austria | | (H) | (H) | (H) | | (L) | (L) | (L) | | (L) | (L) | (L) | | (L) | (L) | (H) | | (L) | (L) | (L) |
| Belgium | | (H) | (H) | (H) | | (L) | (L) | (L) | | (H) | (H) | (H) | | (H) | (H) | (H) | | (L) | (L) | (L) |
| Cyprus | | | (H) | (H) | | | (H) | (H) | | | (L) | (L) | | | (H) | (H) | | | (L) | (L) |
| Denmark | | (H) | (H) | (H) | | (L) | (L) | (L) | | (L) | (L) | (L) | | (L) | (L) | (H) | | (L) | (L) | (L) |
| France | (H) | (H) | (H) | (H) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) |
| Finland | | (H) | (H) | (H) | | (L) | (L) | (L) | | (L) | (L) | (L) | | (L) | (L) | (L) | | (L) | (L) | (L) |
| Germany | | (H) | (H) | (H) | | (L) | (L) | (L) | | (L) | (L) | (L) | | (L) | (L) | (L) | | (L) | (L) | (L) |
| Greece | (H) | (H) | (H) | (H) | (L) | (L) | (L) | (L) | (L) | (L) | (H) | (H) | (L) | (L) | (L) | (L) | (H) | (L) | (H) | (L) |
| Ireland | | (H) | (H) | (H) | | (L) | (L) | (H) | | (H) | (H) | (L) | | (L) | (H) | (H) | | (H) | (L) | (L) |
| Italy | | (H) | (H) | (H) | | (L) | (L) | (L) | | (H) | (H) | (H) | | (L) | (L) | (L) | | (H) | (L) | (L) |
| Netherlands | | (H) | (H) | (H) | | (L) | (L) | (L) | | (L) | (L) | (L) | | (L) | (H) | (H) | | (L) | (L) | (L) |
| Norway | | (H) | (H) | (H) | | (L) | (L) | (L) | | (L) | (L) | (L) | | (L) | (L) | (L) | | (H) | (L) | (H) |
| Portugal | | (H) | (H) | (H) | | (L) | (L) | (L) | | (L) | (L) | (L) | | (L) | (L) | (L) | | (H) | (H) | (L) |
| Sweden | (H) | (H) | (H) | (H) | (L) | (L) | (L) | (L) | (L) | (L) | (H) | (L) | (L) | (L) | (L) | (H) | (L) | (L) | (L) | (L) |
| Switzerland | | | (H) | (H) | | | (L) | (L) | | | (L) | (L) | | | (L) | (H) | | | (L) | (L) |
| Spain | | (H) | (H) | (H) | | (L) | (L) | (H) | | (L) | (L) | (L) | | (L) | (L) | (L) | | (L) | (L) | (L) |
| UK | | (H) | (H) | (H) | | (L) | (L) | (L) | | (L) | (L) | (L) | | (L) | (L) | (L) | | (H) | (L) | (L) |
| **Offshoots** | | | | | | | | | | | | | | | | | | | | |
| Australia | | (H) | (H) | (H) | | (H) | (L) | (L) | | (L) | (L) | (L) | | (L) | (L) | (L) | | (L) | (L) | (L) |
| Canada | | (H) | (H) | (H) | | (L) | (L) | (L) | | (L) | (H) | (H) | | (L) | (L) | (L) | | (L) | (L) | (L) |
| N. Zealand | | | (H) | (H) | | | (L) | (L) | | | (L) | (L) | | | (L) | (L) | | | (L) | (L) |
| USA | (H) | (H) | (H) | (H) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) |
| **Middle East and North Africa** | | | | | | | | | | | | | | | | | | | | |
| Algeria | | (H) | (H) | (H) | | (H) | (H) | (L) | | (L) | (H) | (L) | | (H) | (L) | (L) | | (H) | (H) | (H) |
| Bahrain | | | (H) | | | | (H) | | | | (L) | | | | (H) | | | | (L) | |
| Egypt | | (H) | (H) | (H) | | (H) | (H) | (H) | | (H) | (H) | (H) | | (L) | (L) | (L) | | (H) | (H) | (L) |
| Iran | | (H) | (H) | (H) | | (H) | (H) | (L) | | (L) | (L) | (L) | | (L) | (L) | (L) | | (H) | (H) | (H) |
| Israel | | (H) | (H) | (H) | | (H) | (H) | (H) | | (H) | (H) | (H) | | (L) | (L) | (H) | | (H) | (H) | (L) |
| Jordan | | | | (H) | | | | (H) | | | | (H) | | | | (H) | | | | (H) |
| Morocco | (H) | (H) | (H) | (H) | (H) | (H) | (L) | (L) | (L) | (H) | (H) | (L) | (L) | (L) | (L) | (L) | (H) | (L) | (L) | (L) |
| Tunisia | | (H) | (H) | (H) | | (H) | (H) | (L) | | (L) | (L) | (L) | | (H) | (H) | (H) | | (H) | (L) | (L) |
| Saudi Arabia | | | | (H) | | | | (H) | | | | (L) | | | | (L) | | | | (H) |
| Syria | | (H) | (H) | (H) | | (H) | (H) | (H) | | (H) | (H) | (H) | | (L) | (L) | (L) | | (H) | (H) | (H) |
| United Arab Emirates | | | | (H) | | | | (H) | | | | (L) | | | | (H) | | | | (H) |
| Yemen | | | | (H) | | | | (H) | | | | (L) | | | | (L) | | | | (H) |

## Table 4.6 continued

| Threshold | *Initial Income* | | | | *Pop. Growth* | | | | *Debt* | | | | *Openness* | | | | *Inf. Vol.* | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (1) | (2) | (3) | (4) | (1) | (2) | (3) | (4) | (1) | (2) | (3) | (4) | (1) | (2) | (3) | (4) |
| **Sub-Sahara** | | | | | | | | | | | | | | | | | | | | |
| Benin | | (H) | (H) | (H) | | (H) | (H) | (H) | | (L) | (L) | (L) | | (L) | (L) | (L) | | (H) | (H) | (L) |
| Botswana | | (H) | (H) | (H) | | (H) | (H) | (H) | | (L) | (L) | (L) | | (H) | (H) | (H) | | (H) | (H) | (H) |
| Burundi | (L) | (L) | (L) | (L) | (H) | (H) | (H) | (H) | (L) | (L) | (H) | (H) | (L) | (L) | (L) | (L) | (H) | (H) | (H) | (H) |
| Cameroon | | (H) | (H) | (H) | | (H) | (H) | (H) | | (L) | (H) | (L) | | (L) | (L) | (L) | | (H) | (H) | (L) |
| Cen. Afr. Rep. | | (H) | (L) | (L) | | (H) | (H) | (H) | | (L) | (H) | (H) | | (L) | (L) | (L) | | (H) | (H) | (L) |
| Cote d‘Ivoire | | (H) | (H) | (H) | | (H) | (H) | (H) | | (H) | (H) | (H) | | (L) | (L) | (H) | | (H) | (H) | (L) |
| Congo, Dem. Rep. | | (L) | (L) | | | (H) | (H) | | | (L) | (H) | | | (L) | (H) | | | (H) | (H) | |
| Congo, Rep. | | (H) | (H) | (H) | | (H) | (H) | (H) | | (H) | (H) | (H) | | (H) | (H) | (H) | | (H) | (H) | (H) |
| Gabon | | (H) | (H) | (H) | | (H) | (H) | (H) | | (L) | (H) | (L) | | (H) | (H) | (H) | | (H) | (H) | (H) |
| Gambia, The | | (H) | (H) | (L) | | (H) | (H) | (H) | | (H) | (H) | (H) | | (L) | (L) | (L) | | (H) | (L) | (H) |
| Ghana | (H) | (L) | (H) | (H) | (H) | (H) | (H) | (H) | (L) | (L) | (L) | (L) | (H) | (L) | (H) | (H) | (H) | (H) | (H) | (H) |
| Kenya | (H) | (H) | (H) | (H) | (H) | (H) | (H) | (H) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (H) | (L) | (H) | (L) |
| Lesotho | | (L) | (H) | (H) | | (H) | (L) | (L) | | (L) | (H) | (L) | | (H) | (H) | (H) | | (H) | (L) | (H) |
| Liberia | | (H) | | | | (L) | | | | (H) | | | | (H) | | | | (H) | | |
| Mali | | (L) | (L) | (L) | | (H) | (H) | (H) | | (H) | (H) | (L) | | (L) | (L) | (L) | | (H) | (H) | (H) |
| Malawi | | (H) | (L) | (L) | | (H) | (H) | (H) | | (H) | (H) | (H) | | (L) | (L) | (L) | | (H) | (H) | (H) |
| Mauritania | | (H) | (H) | (H) | | (H) | (H) | (H) | | (H) | (H) | (H) | | (H) | (H) | (H) | | (L) | (H) | (H) |
| Mauritius | | | (H) | (H) | | | (L) | (L) | | | (L) | (L) | | | (H) | (H) | | | (L) | (L) |
| Mozambique | | | | (L) | | | | (H) | | | | (H) | | | | (L) | | | | (L) |
| Namibia | | | | (H) | | | | (L) | | | | (L) | | | | (H) | | | | (H) |
| Niger | | (L) | (L) | (L) | | (H) | (H) | (H) | | (L) | (L) | (L) | | (L) | (L) | (L) | | (H) | (H) | (L) |
| Rwanda | | (L) | (L) | (L) | | (H) | (H) | (H) | | (L) | (H) | (L) | | (H) | (L) | (L) | | (H) | (H) | (H) |
| South Africa | (H) | (H) | (H) | (H) | (H) | (H) | (H) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (H) | (H) | (L) | (L) |
| Senegal | | (H) | (H) | (H) | | (H) | (H) | (H) | | (L) | (H) | (L) | | (L) | (L) | (L) | | (H) | (H) | (L) |
| Sudan | | | | (H) | | | | (H) | | | | (H) | | | | (L) | | | | (H) |
| Sierra Leone | | (H) | (H) | (L) | | (H) | (L) | (H) | | (H) | (H) | (H) | | (L) | (L) | (L) | | (H) | (H) | (H) |
| Swaziland | | (H) | (H) | (H) | | (H) | (H) | (H) | | (L) | (L) | (L) | | (H) | (H) | (H) | | (H) | (H) | (L) |
| Zambia | | (H) | (H) | (H) | | (H) | (H) | (H) | | (H) | (H) | (H) | | (L) | (L) | (L) | | (H) | (H) | (H) |
| Zimbabwe | | (L) | (L) | (L) | | (H) | (H) | (L) | | (L) | (L) | (L) | | (L) | (L) | (H) | | (H) | (H) | (H) |
| Uganda | | | (L) | (L) | | | (H) | (H) | | | (H) | (L) | | | (L) | (L) | | | (H) | (H) |
| Tanzania | | | | (L) | | | | (H) | | | | (L) | | | | (L) | | | | (L) |
| Togo | | (H) | (H) | (H) | | (H) | (H) | (H) | | (H) | (H) | (H) | | (L) | (L) | (H) | | (H) | (H) | (H) |

## Table 4.6 continued

| Threshold | Initial Income | | | | Pop. Growth | | | | Debt | | | | Openness | | | | Inf. Vol. | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (1) | (2) | (3) | (4) | (1) | (2) | (3) | (4) | (1) | (2) | (3) | (4) | (1) | (2) | (3) | (4) |
| **Latin and Caribbean** | | | | | | | | | | | | | | | | | | | | |
| Argentina | | (H) | (H) | (H) | | (H) | (L) | (L) | | (L) | (L) | (H) | | (L) | (L) | (L) | | (H) | (H) | (H) |
| Bolivia | | (H) | (H) | (H) | | (H) | (H) | (H) | | (H) | (H) | (L) | | (L) | (L) | (L) | | (H) | (H) | (H) |
| Brazil | | (H) | (H) | (H) | | (H) | (H) | (L) | | (L) | (L) | (L) | | (L) | (L) | (L) | | (H) | (H) | (L) |
| Chile | | (H) | (H) | (H) | | (H) | (H) | (L) | | (H) | (L) | (L) | | (L) | (L) | (L) | | (H) | (H) | (L) |
| Colombia | (H) | (H) | (H) | (H) | (H) | (H) | (H) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (H) | (L) | (H) | (H) |
| Costa Rica | (H) | (H) | (H) | (H) | (H) | (H) | (H) | (H) | (L) | (H) | (L) | (L) | (L) | (L) | (L) | (H) | (H) | (H) | (H) | (L) |
| Dom. Rep | | (H) | (H) | (H) | | (H) | (H) | (H) | | (L) | (L) | (L) | | (L) | (L) | (L) | | (H) | (H) | (H) |
| Ecuador | (H) | (H) | (H) | (H) | (H) | (H) | (H) | (H) | (L) | (L) | (H) | (L) | (L) | (L) | (L) | (L) | (H) | (H) | (H) | (H) |
| El Salvador | (H) | (H) | (H) | (H) | (H) | (L) | (H) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (H) | (H) | (L) | (L) |
| Guyana | | (H) | (H) | (H) | | (L) | (L) | (L) | | (H) | (H) | (H) | | (H) | (H) | (H) | | (H) | (H) | (H) |
| Guatemala | (H) | (H) | (H) | (H) | (H) | (H) | (H) | (H) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (H) | (H) | (H) | (L) |
| Haiti | | | | (H) | | | | (H) | | | | (L) | | | | (L) | | | | (H) |
| Honduras | (H) | (H) | (H) | (H) | (H) | (H) | (H) | (H) | (L) | (L) | (H) | (L) | (H) | (H) | (H) | (H) | (H) | (L) | (H) | (H) |
| Jamaica | | (H) | (H) | (H) | | (L) | (L) | (L) | | (H) | (H) | (H) | | (H) | (H) | (H) | | (H) | (H) | (L) |
| Mexico | (H) | (H) | (H) | (H) | (H) | (H) | (H) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (H) | (H) | (H) | (H) |
| Nicaragua | | (H) | (H) | (H) | | (H) | (H) | (H) | | (H) | (H) | (H) | | (L) | (L) | (H) | | (H) | (H) | (L) |
| Panama | (H) | (H) | (H) | (H) | (H) | (H) | (H) | (H) | (L) | (H) | (L) | (L) | (H) | (H) | (H) | (H) | (L) | (H) | (H) | (L) |
| Paraguay | | (H) | (H) | (H) | | (H) | (H) | (H) | | (L) | (L) | (L) | | (L) | (H) | (H) | | (H) | (H) | (H) |
| Peru | | (H) | (H) | (H) | | (H) | (H) | (L) | | (L) | (L) | (L) | | (L) | (L) | (L) | | (H) | (H) | (L) |
| Venezuela | (H) | (H) | (H) | (H) | (H) | (H) | (H) | (H) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (H) | (H) | (H) | (H) |
| Trin.& Tob. | (H) | (H) | (H) | (H) | (L) | (H) | (L) | (L) | (L) | (L) | (L) | (L) | (H) | (H) | (H) | (H) | (H) | (H) | (H) | (H) |
| Uruguay | | (H) | (H) | (H) | | (L) | (L) | (L) | | (L) | (L) | (L) | | (L) | (L) | (L) | | (H) | (H) | (H) |
| **South Asia** | | | | | | | | | | | | | | | | | | | | |
| Bangladesh | | (L) | (L) | (H) | | (H) | (H) | (H) | | (L) | (L) | (L) | | (L) | (L) | (L) | | (L) | (L) | (L) |
| India | (H) | (H) | (H) | (H) | (H) | (H) | (H) | (H) | (L) | (L) | (L) | (H) | (L) | (L) | (L) | (L) | (H) | (L) | (L) | (L) |
| Nepal | | (L) | (H) | (H) | | (H) | (H) | (H) | | (L) | (L) | (L) | | (L) | (L) | (L) | | (L) | (L) | (H) |
| Pakistan | (H) | (H) | (H) | (H) | (H) | (H) | (H) | (H) | (L) | (L) | (H) | (L) | (L) | (L) | (L) | (L) | (H) | (L) | (L) | (H) |
| Sri Lanka | (H) | (H) | (H) | (H) | (H) | (H) | (L) | (L) | (L) | (H) | (H) | (H) | (L) | (L) | (L) | (L) | (H) | (H) | (L) | (L) |

## Table 4.6 continued

| Threshold | *Initial Income* (1) | (2) | (3) | (4) | *Pop. Growth* (1) | (2) | (3) | (4) | *Debt* (1) | (2) | (3) | (4) | *Openness* (1) | (2) | (3) | (4) | *Inf. Vol.* (1) | (2) | (3) | (4) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **East Asia and Pacific** | | | | | | | | | | | | | | | | | | | | |
| Cambodia | | | | (H) | | | | (H) | | | | (L) | | | | (H) | | | | (H) |
| China | | | (H) | (H) | | | (L) | (L) | | | (L) | (L) | | | (L) | (L) | | | (H) | (L) |
| Fiji | | (H) | (H) | (H) | | (H) | (L) | (L) | | (L) | (L) | (L) | | (H) | (H) | (H) | | (H) | (L) | (L) |
| Indonesia | | (H) | (H) | (H) | | (H) | (H) | (L) | | (L) | (L) | (L) | | (L) | (L) | (L) | | (H) | (L) | (H) |
| Japan | | (H) | (H) | (H) | | (L) | (L) | (L) | | (L) | (H) | (H) | | (L) | (L) | (L) | | (L) | (L) | (L) |
| Korea Rep. | (H) | (H) | (H) | (H) | (H) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (H) | (H) | (L) | (L) |
| Laos | | | (H) | (H) | | | (H) | (H) | | | (H) | (H) | | | (L) | (L) | | | (H) | (H) |
| Malaysia | (H) | (H) | (H) | (H) | (H) | (H) | (H) | (H) | (L) | (H) | (L) | (L) | (H) | (H) | (H) | (H) | (H) | (H) | (L) | (H) |
| Myanmar | | | | | | | | | | | | | | | | | | | | |
| Mongolia | | | | (H) | | | | (H) | | | | (L) | | | | (H) | | | | (H) |
| Papua New Guinea | | (H) | (H) | (H) | | (H) | (H) | (H) | | (L) | (L) | (L) | | (L) | (H) | (H) | | (H) | (H) | (H) |
| Philippines | (H) | (H) | (H) | (H) | (H) | (H) | (H) | (H) | (L) | (L) | (L) | (L) | (L) | (L) | (H) | (H) | (H) | (H) | (L) | (L) |
| Singapore | (H) | (H) | (H) | (H) | (H) | (H) | (H) | (H) | (L) | (H) | (H) | (H) | (H) | (H) | (H) | (H) | (H) | (L) | (L) | (L) |
| Thailand | (H) | (H) | (H) | (H) | (H) | (H) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (H) | (H) | (H) | (L) | (L) | (L) |
| Vietnam | | | | (H) | | | | (L) | | | | (L) | | | | (H) | | | | (H) |
| **East Europe and Central Asia** | | | | | | | | | | | | | | | | | | | | |
| Albania | | | | (H) | | | | (L) | | | | (L) | | | | (L) | | | | (L) |
| Armenia | | | | (H) | | | | (L) | | | | (L) | | | | (L) | | | | (L) |
| Bulgaria | | | | (H) | | | | (L) | | | | (L) | | | | (H) | | | | (L) |
| Croatia | | | | (H) | | | | (L) | | | | (L) | | | | (H) | | | | (L) |
| Czech Republic | | | | (H) | | | | (L) | | | | (L) | | | | (H) | | | | (L) |
| Estonia | | | | (H) | | | | (L) | | | | (L) | | | | (H) | | | | (L) |
| Hungary | | | (H) | (H) | | | (L) | (L) | | | (H) | (L) | | | (L) | (H) | | | (H) | (L) |
| Kazakhstan | | | | (H) | | | | (L) | | | | (L) | | | | (H) | | | | (H) |
| Kyrgyzstan | | | | (H) | | | | (L) | | | | (H) | | | | (H) | | | | (H) |
| Latvia | | | | (H) | | | | (L) | | | | (L) | | | | (H) | | | | (H) |
| Lithuania | | | | (H) | | | | (L) | | | | (L) | | | | (H) | | | | (H) |
| Moldova | | | | (H) | | | | (L) | | | | (L) | | | | (H) | | | | (H) |
| Poland | | | | (H) | | | | (L) | | | | (L) | | | | (L) | | | | (L) |
| Romania | | | | (H) | | | | (L) | | | | (L) | | | | (L) | | | | (H) |
| Russia | | | | (H) | | | | (L) | | | | (L) | | | | (L) | | | | (H) |
| Slovak Republic | | | | (H) | | | | (L) | | | | (L) | | | | (H) | | | | (L) |
| Slovenia | | | | (H) | | | | (L) | | | | (L) | | | | (H) | | | | (L) |
| Tajikistan | | | | (H) | | | | (H) | | | | (L) | | | | (H) | | | | (H) |
| Turkey | (H) | (H) | (H) | (H) | (H) | (H) | (H) | (H) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (L) | (H) | (H) | (H) | (H) |
| Ukraine | | | | (H) | | | | (L) | | | | (L) | | | | (H) | | | | (H) |

## Table 4.7: STR-BMA for the significant threshold variables

This table presents modeling averaging results of equation (4.2.6) using LS-BMA. The posterior inclusion probability (PIP) of a variable is the sum of the posterior probabilities models that include that variable. The posterior mean (PM) is the average of the coefficient estimates (COEF) of individual models weighted by posterior probability. The posterior standard error (PSE) is the BMA estimate for the standard error (SE) taking model uncertainty into account. Constant and trend were always included in the regressions. In the model space we include threshold effects based only on significant threshold variables. For the threshold variables of Openness and Inflation Volatility we use their logarithmic transformation.

| | Model averaging | | | Posterior mode | | Largest | |
|---|---|---|---|---|---|---|---|
| | PIP | PM | PSE | COEF | SE | COEF | SE |
| Executive Constraints | 0.9870 | -0.0076 | 0.0019 | -0.0082 | 0.0013 | -0.0039 | 0.0019 |
| Executive Constraints $\times I(\text{Debt} < \widehat{\gamma}_1)$ | 1.0000 | 0.0085 | 0.0017 | 0.0090 | 0.0015 | 0.0064 | 0.0017 |
| Executive Constraints $\times I(\text{Population Growth} < \widehat{\gamma}_2)$ | 0.1390 | -0.0005 | 0.0015 | - | - | -0.0028 | 0.0026 |
| Executive Constraints $\times I(\text{Inflation Volatility} < \widehat{\gamma}_3)$ | 0.1580 | 0.0003 | 0.0012 | - | - | 0.0007 | 0.0018 |
| Executive Constraints $\times I(\text{Openness} < \widehat{\gamma}_4)$ | 0.3340 | -0.0006 | 0.0011 | - | - | -0.0028 | 0.0016 |
| Executive Constraints $\times I(\text{Initial Income} < \widehat{\gamma}_5)$ | 0.4293 | -0.0032 | 0.0043 | -0.0077 | 0.0031 | -0.0059 | 0.0044 |
| Population Growth | 0.0600 | -0.0010 | 0.0077 | - | - | -0.0387 | 0.0194 |
| Population Growth $\times I(\text{Debt} < \widehat{\gamma}_1)$ | 0.9927 | 0.0187 | 0.0058 | 0.0181 | 0.0028 | 0.0324 | 0.0133 |
| Population Growth $\times I(\text{Population Growth} < \widehat{\gamma}_2)$ | 0.9897 | -0.0377 | 0.0102 | -0.0390 | 0.0077 | -0.0298 | 0.0197 |
| Population Growth $\times I(\text{Inflation Volatility} < \widehat{\gamma}_3)$ | 0.8510 | 0.0089 | 0.0095 | 0.0051 | 0.0012 | 0.0294 | 0.0097 |
| Population Growth $\times I(\text{Openness} < \widehat{\gamma}_4)$ | 0.5097 | 0.0035 | 0.0064 | 0.0037 | 0.0011 | 0.0302 | 0.0151 |
| Population Growth $\times I(\text{Initial Income} < \widehat{\gamma}_5)$ | 0.8820 | -0.0140 | 0.0116 | -0.0167 | 0.0039 | -0.0243 | 0.0402 |
| Life Expectancy | 0.2890 | -0.0002 | 0.0004 | - | - | -0.0020 | 0.0007 |
| Life Expectancy $\times I(\text{Debt} < \widehat{\gamma}_1)$ | 0.1700 | 0.0001 | 0.0003 | - | - | 0.0008 | 0.0005 |
| Life Expectancy $\times I(\text{Population Growth} < \widehat{\gamma}_2)$ | 0.9643 | -0.0014 | 0.0005 | -0.0015 | 0.0003 | -0.0010 | 0.0007 |
| Life Expectancy $\times I(\text{Inflation Volatility} < \widehat{\gamma}_3)$ | 0.3340 | 0.0001 | 0.0004 | - | - | 0.0010 | 0.0004 |
| Life Expectancy $\times I(\text{Openness} < \widehat{\gamma}_4)$ | 0.2930 | 0.0000 | 0.0002 | - | - | 0.0013 | 0.0006 |
| Life Expectancy $\times I(\text{Initial Income} < \widehat{\gamma}_5)$ | 0.2820 | -0.0002 | 0.0006 | - | - | -0.0006 | 0.0020 |
| $\text{IMR}_{\text{Debt}}$ | 0.1463 | -0.0005 | 0.0016 | - | - | 0.0012 | 0.0033 |
| $\text{IMR}_{\text{Population Growth}}$ | 0.0750 | 0.0001 | 0.0010 | - | - | 0.0053 | 0.0050 |
| $\text{IMR}_{\text{Inflation Volatility}}$ | 0.0653 | -0.0002 | 0.0012 | - | - | -0.0044 | 0.0031 |
| $\text{IMR}_{\text{Openness}}$ | 0.1060 | -0.0005 | 0.0016 | - | - | -0.0041 | 0.0042 |
| $\text{IMR}_{\text{Initial Income}}$ | 0.0740 | 0.0004 | 0.0018 | - | - | 0.0036 | 0.0041 |

## Table 4.8: STR-BMA for all threshold variables

This table presents modeling averaging results of equation (4.2.6) using LS-BMA. The posterior inclusion probability (PIP) of a variable is the sum of the posterior probabilities models that include that variable. The posterior mean (PM) is the average of the coefficient estimates (COEF) of individual models weighted by posterior probability. The posterior standard error (PSE) is the BMA estimate for the standard error (SE) taking model uncertainty into account. Constant and trend were always included in the regressions. In the model space we include the threshold effects from all possible threshold variable as well as the corresponding inverse Mills ratio terms. For the threshold variables of Openness and Inflation Volatility we use their logarithmic transformation.

| | Model averaging | | | Posterior mode | | Largest | |
|---|---|---|---|---|---|---|---|
| | PIP | PM | PSE | COEF | SE | COEF | SE |
| Executive Constraints | 0.8013 | -0.0048 | 0.0030 | -0.0050 | 0.0014 | -0.0013 | 0.0022 |
| Executive Constraints $\times I(\text{Debt} < \widehat{\gamma}_1)$ | 1.0000 | 0.0087 | 0.0021 | 0.0079 | 0.0015 | 0.0072 | 0.0019 |
| Executive Constraints $\times I(\text{Executive Constraints} < \widehat{\gamma}_2$ | 0.2427 | 0.0035 | 0.0099 | - | - | 0.0504 | 0.0445 |
| Executive Constraints $\times I(\text{Gover} < \widehat{\gamma}_3)$ | 0.1863 | -0.0004 | 0.0013 | - | - | -0.0026 | 0.0017 |
| Executive Constraints $\times I(\text{Population Growth} < \widehat{\gamma}_4)$ | 0.1437 | -0.0004 | 0.0014 | - | - | -0.0038 | 0.0023 |
| Executive Constraints $\times I(\text{Inflation} < \widehat{\gamma}_5)$ | 0.5503 | -0.0030 | 0.0032 | - | - | -0.0026 | 0.0032 |
| Executive Constraints $\times I(\text{Inflation Volatility} < \widehat{\gamma}_6)$ | 0.2980 | 0.0000 | 0.0016 | - | - | 0.0001 | 0.0023 |
| Executive Constraints $\times I(\text{Investments} < \widehat{\gamma}_7$ | 0.4563 | -0.0029 | 0.0037 | -0.0072 | 0.0025 | -0.0061 | 0.0030 |
| Executive Constraints $\times I(\text{Life} < \widehat{\gamma}_8)$ | 0.6990 | -0.0027 | 0.0022 | -0.0036 | 0.0012 | -0.0030 | 0.0023 |
| Executive Constraints $\times I(\text{Openness} < \widehat{\gamma}_9)$ | 0.0793 | -0.0002 | 0.0008 | - | - | -0.0009 | 0.0018 |
| Executive Constraints $\times I(\text{Schooling} < \widehat{\gamma}_{10})$ | 0.0170 | 0.0000 | 0.0002 | - | - | -0.0015 | 0.0045 |
| Executive Constraints $\times I(\text{Initial Income} < \widehat{\gamma}_{11})$ | 0.0920 | 0.0001 | 0.0007 | - | - | -0.0030 | 0.0050 |
| Population Growth | 0.2353 | -0.0068 | 0.0148 | - | - | -0.0410 | 0.0202 |
| Population Growth $\times I(\text{Debt} < \widehat{\gamma}_1)$ | 0.5473 | 0.0091 | 0.0091 | 0.0158 | 0.0027 | 0.0239 | 0.0102 |
| Population Growth $\times I(\text{Executive Constraints} < \widehat{\gamma}_2)$ | 0.3713 | -0.0020 | 0.0051 | - | - | 0.0308 | 0.0366 |
| Population Growth $\times I(\text{Gov} < \widehat{\gamma}_3)$ | 0.1570 | 0.0009 | 0.0043 | - | - | 0.0154 | 0.0112 |
| Population Growth $\times I(\text{Population Growth} < \widehat{\gamma}_4)$ | 0.3480 | -0.0037 | 0.0079 | - | - | -0.0290 | 0.0149 |
| Population Growth $\times I(\text{Inflation} < \widehat{\gamma}_5)$ | 0.7820 | -0.0173 | 0.0169 | -0.0333 | 0.0094 | -0.0412 | 0.0193 |

Table continued on next page ...

Table 4.8 continued

| | Model averaging | | | Posterior mode | | Largest | |
|---|---|---|---|---|---|---|---|
| | PIP | PM | PSE | COEF | SE | COEF | SE |
| Population Growth $\times I$(Inflation Volatility $< \widehat{\gamma}_6$) | 0.8490 | 0.0108 | 0.0128 | 0.0059 | 0.0012 | 0.0379 | 0.0115 |
| Population Growth $\times I$(Investments $< \widehat{\gamma}_7$) | 0.9793 | -0.0462 | 0.0143 | -0.0476 | 0.0090 | -0.0227 | 0.0210 |
| Population Growth $\times I$(Life $< \widehat{\gamma}_8$) | 0.1490 | 0.0000 | 0.0032 | - | - | 0.0042 | 0.0263 |
| Population Growth $\times I$(Openness $< \widehat{\gamma}_9$) | 0.8367 | 0.0080 | 0.0096 | 0.0045 | 0.0010 | 0.0310 | 0.0130 |
| Population Growth $\times I$(Schooling $< \widehat{\gamma}_{10}$) | 0.0523 | -0.0006 | 0.0055 | - | - | -0.0576 | 0.0456 |
| Population Growth $\times I$(Initial Income $< \widehat{\gamma}_{11}$) | 0.1490 | -0.0005 | 0.0023 | - | - | 0.0118 | 0.0286 |
| Life Expectancy | 0.6703 | -0.0009 | 0.0008 | -0.0009 | 0.0003 | -0.0024 | 0.0008 |
| Life Expectancy $\times I$(Debt $< \widehat{\gamma}_1$) | 0.5870 | -0.0003 | 0.0005 | - | - | 0.0004 | 0.0005 |
| Life Expectancy $\times I$(Executive Constraints $< \widehat{\gamma}_2$) | 0.2657 | 0.0001 | 0.0002 | - | - | 0.0006 | 0.0013 |
| Life Expectancy $\times I$(Government $< \widehat{\gamma}_3$) | 0.2660 | 0.0001 | 0.0002 | - | - | 0.0009 | 0.0005 |
| Life Expectancy $\times I$(Population Growth $< \widehat{\gamma}_4$) | 0.2673 | -0.0001 | 0.0003 | - | - | -0.0009 | 0.0006 |
| Life Expectancy $\times I$(Inflation $< \widehat{\gamma}_5$) | 0.2977 | -0.0004 | 0.0008 | -0.0012 | 0.0004 | -0.0013 | 0.0008 |
| Life Expectancy $\times I$(Inflation Volatility $< \widehat{\gamma}_6$) | 0.2033 | 0.0002 | 0.0005 | - | - | 0.0014 | 0.0005 |
| Life Expectancy $\times I$(Investments $< \widehat{\gamma}_7$) | 0.9883 | -0.0021 | 0.0007 | -0.0019 | 0.0005 | -0.0008 | 0.0010 |
| Life Expectancy $\times I$(Life $< \widehat{\gamma}_8$) | 0.0950 | 0.0000 | 0.0001 | - | - | 0.0003 | 0.0012 |
| Life Expectancy $\times I$(Openness $< \widehat{\gamma}_9$) | 0.3903 | 0.0002 | 0.0004 | - | - | 0.0012 | 0.0006 |
| Life Expectancy $\times I$(Schooling $< \widehat{\gamma_{10}}$) | 0.0343 | 0.0000 | 0.0003 | - | - | -0.0028 | 0.0022 |
| Life Expectancy $\times I$(Initial Income $< \widehat{\gamma_{11}}$) | 0.1170 | 0.0000 | 0.0001 | - | - | 0.0009 | 0.0014 |

Table continued on next page ...

Table 4.8 continued

| | Model averaging | | | Posterior mode | | Largest | |
|---|---|---|---|---|---|---|---|
| | PIP | PM | PSE | COEF | SE | COEF | SE |
| $IMR_{Debt}$ | 0.0617 | -0.0001 | 0.0008 | - | - | 0.0014 | 0.0037 |
| $IMR_{Executive\ Constraints}$ | 0.0730 | 0.0002 | 0.0011 | - | - | -0.0026 | 0.0041 |
| $IMR_{Government}$ | 0.1143 | 0.0002 | 0.0012 | - | - | -0.0015 | 0.0058 |
| $IMR_{Population\ Growth}$ | 0.0650 | 0.0001 | 0.0009 | - | - | 0.0029 | 0.0055 |
| $IMR_{Inflation}$ | 0.0577 | 0.0000 | 0.0006 | - | - | -0.0008 | 0.0024 |
| $IMR_{Inflation\ Volatility}$ | 0.0633 | -0.0001 | 0.0008 | - | - | -0.0032 | 0.0034 |
| $IMR_{Investments}$ | 0.0260 | 0.0000 | 0.0005 | - | - | -0.0037 | 0.0046 |
| $IMR_{Life}$ | 0.0523 | 0.0001 | 0.0007 | - | - | 0.0057 | 0.0047 |
| $IMR_{Openness}$ | 0.0433 | -0.0001 | 0.0008 | - | - | -0.0035 | 0.0044 |
| $IMR_{Schooling}$ | 0.0873 | 0.0005 | 0.0021 | - | - | 0.0091 | 0.0097 |
| $IMR_{Initial\ Income}$ | 0.1943 | 0.0010 | 0.0028 | - | - | 0.0019 | 0.0098 |

## Table B1: Data Appendix

| Variable | Description |
|---|---|
| Time trend | Time trend variable for the periods 1970-79, 1980-89, 1990-99 and 2000-2009. |
| Growth volatility | Standard deviation of the growth rate of real per capita GDP for the periods 1970-79, 1980-89, 1990-99 and 2000-2009. Source: PWT 7.0. |
| Initial Income | Logarithm of per capita GDP in chain series at 1970, 1980, 1990, 2000. Source: PWT 7.0. |
| Lag Initial Income | Logarithm of per capita GDP in chain series at 1965, 1975, 1985 and 1995. Source: PWT 7.0. |
| Population Growth Rates | Logarithm of average population growth rates plus 0.05 for the periods 1970-79, 1980-89, 1990-99 and 2000-2009. Source: PWT 7.0. |
| Lag Population Growth Rates | Logarithm of average population growth rates plus 0.05 for the periods 1965-69, 1975-79, 1985-89 and 1995-1999. Source: PWT 7.0. |
| Investment | Logarithm of average ratios over each period of investment to GDP for the periods 1970-79, 1980-89, 1990-99 and 2000-2009. Source: PWT 7.0. |
| Lag Investement | Logarithm of average ratios over each period of investment to GDP for the periods 1965-69, 1975-79, 1985-89 and 1995-1999. Source: PWT 7.0. |
| Schooling | Logarithm of average years of male secondary and tetriary school attainment (25+) in 1970, 1980, 1990, and 1999. Source: Barro and Lee (2000). |
| Lag Schooling | Logarithm of average years of male secondary and tetriary school attainment (25+) in 1965, 1975, 1985 and 1995. Source: Barro and Lee (2000). |
| Debt | Logarithm of public debt to GDP for the periods 1970-79, 1980-89, 1990-99 and 2000-2009. Source: IMF, Debt Database Fall 2011 Vintage |
| Lag Debt | Logarithm of public debt to GDP for the periods 1965-69, 1975-79, 1985-89 and 1995-1999. Source: IMF, Debt Database Fall 2011 Vintage |
| Government | Log of average ratios for each period of government consumption (net of outlays on defense and education) to GDP for the periods 1970-79, 1980-89, 1990-99 and 2000-2009. Source: PWT 7.0 |
| Lag Government | Log of average ratios for each period of government consumption (net of outlays on defense and education) to GDP for the periods 1965-69, 1975-79, 1985-89 and 1995-1999. Source: PWT 7.0 |

Table continued on next page ...

Table B1 continued

| **Variable** | **Description** |
|---|---|
| Inflation | Log average inflation for the periods 1970-79, 1980-89, 1990-99 and 2000-2009. Source: Worldbank |
| Lag Inflation | Log average inflation for the periods 1965-69, 1975-79, 1985-89 and 1995-1999. Source: Worldbank |
| Inflation Volatility | Standard deviation of inflation for the periods 1970-79, 1980-89, 1990-99 and 2000-2009. Source: Worldbank |
| Lag Inflation Volatility | Standard deviation of inflation for the periods 1965-69, 1975-79, 1985-89 and 1995-1999. Source: Worldbank |
| Openness | Average ratios for each period of exports plus imports to GDP for the periods 1970-79, 1980-89, 1990-99 and 2000-2009. Source: PWT 7.0. |
| Lag of Openness | Average ratios for each period of exports plus imports to GDP for the periods 1965-69, 1975-79, 1985-89 and 1995-1999. Source: PWT 7.0. |
| Life Expectancy | Log of average life expectancy for the periods 1970-79, 1980-89, 1990-99 and 2000-2009. Source: World Bank |
| Lag Life Expectancy | Log of average life expectancy for the periods 1965-69, 1975-79, 1985-89 and 1995-1999. Source: World Bank |
| Executive Constraints | A measure of the extent of institutionalized constraints on the decision making powers of chief executives. This variable ranges from one to seven where higher values equal a greater extent of institutionalized constraints on the power of chief executives |
| Lag Executive Constraints | A measure of the extent of institutionalized constraints on the decision making powers of chief executives. This variable ranges from one to seven where higher values equal a greater extent of institutionalized constraints on the power of chief executives |
| Tropics | Percentage of land area classified as tropical and subtropical via the in Koeppen-Geiger system. Source: The Center for International Development at Harvard University |
| LCR100KM | Percentage of a country's land area within 100km of an ice- free coast. Source: The Center for International Development at Harvard University. |

# References


Acemoglu, D., and S. Johnson, 2005, Unbundling Institutions, Journal of Political Economy 113, 949–995.

Acemoglu, D., S. Johnson, and J. Robinson, 2001, The Colonial Origins of Comparative Development : An Empirical Investigation, American Economic Review 91, 1369–1401.

Acemoglu, D., and J. Robinson, 2006, Economic Backwardness in Political Perspective, American Political Science Review 100, 115–131.

Acemoglu, D., S. Johnson, and J. Robinson, and Y. Thaicharoen**,** 2003, Institutional causes, macroeconomic symptoms: volatility, crises and growth, Journal of Monetary Economics 50, 49–123.

Acemoglu, D., and F. Zilibotti, 1997, Was prometheus unbound by chance? risk, diversification and growth., Journal of Political Economy 105, 1167–1200.

Alesina, A., R. Baqir, and W. Easterly, 1999, Public goods and ethnic divisions, Quarterly Journal of Economics 114, 1243–1284.

Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg, 2003, Fractionalization, Journal of Economic Growth 8, 155–194.

Bai, J. and P. Perron, 1998, Estimating and Testing Linear Models with Multiple Structural Changes, Econometrica, 66, 47–78.

Banerjee, A., and L. Iyer, 2005, History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India, American Economic Review 95, 1190–1213.

Banks, A, 1999, Cross-National Time-Series Data Archive (Binghamton, New York: Databanks International).

Barro, B. J., and J.-W. Lee, 2010, A New Data Set of Educational Attainment in the World, 1950 to 2010, NBER Working Papers, 15902.

Bekaert, G., C.R. Harvey, and C. Lundblad, 2006, Growth volatility and financial liberalization, Journal of International Money and Finance 92, 370–403.

Bourguignon, F., and C. Morrisson, 2002, Inequality among world citizens: 1890-1992, American Economic Review 92, 1190–1213.

Brock, W., and S. D. Durlauf, 2001, Growth Empirics and Reality, World Bank Economic Review 15, 229–272.

Caner, M., and B. Hansen, 2004, Instrumental variable estimation of a threshold model, Econometric Theory 20, 813–843.

Ciccone, A., and M. Jarocinski, 2010, Determinants of Economic Growth: Will Data Tell?, American Economic Journal: Macroeconomics 2, 222–246.

Comin, D., and B. Hobijn, 2004, Cross-Country Technological Adoption: Making the Theories Face the Facts, Journal of Monetary Economics 51, 39–83.

Cox, D. R., 1972, Regression Models and Life Tables, Journal of the Royal Statistical Society Series B 34, 187–220.

Cuaresma, C. J., and G. Doppelhofer, 2007, Nonlinearities in cross-country growth regressions: A bayesian averaging of thresholds (bat) approach, Journal of Macroeconomics 29, 541–554.

Davies, R. B., 1977, Hypothesis testing when a nuisance parameter is present only under the alternative, Biometrika 64, 247–254.

Desmet, K., and P. Parente, 2009, The Evolution of Markets and the Revolution of Industry: A Unified Theory of Growth, CEPR Discussion Paper 7290.

Dickey, D. A and W. A. Fuller, 1979, Distribution of the estimators for autoregressive time series with a unit root, Journal of the American Statistical Association, 74, 427-431.

Dickey, D. A and W. A. Fuller, 1981, Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root, Econometrica, 49, 1057-1072.

Di Giovanni, J., and A. A. Levchenko, 2009, Trade openness and volatility, The Review of Economics and Statistics 91, 558–585.

Draper, D., 1995, The geography of output volatility, Journal of the Royal Statistical Society, Series B 57, 45–97.

Durlauf, S., and P. Johnson, 1995, Multiple regimes and cross-country growth behavior, Journal of Applied Econometrics 10, 365–384.

Durlauf, S., A. Kourtellos, and A. Minkin, 2001, The local solow growth model, European Economic Review 15, 928–940.

Durlauf, S. N., P. Johnson and J. R. W. Temple, 2005, Growth Econometrics, Handbook of Economic Growth, 1, 1 (8), 555-677.

Durlauf, S. N., A. Kourtellos, and C. M. Tan, 2008, Are Any Growth Theories Robust?, Economic Journal 118, 329–346.

Durlauf, S.N., A. Kourtellos, and C.M. Tan, 2011, Is God in the Details? A Reexamination of the Role of Religion in Economic Growth, Journal of Applied Econometrics, forthcoming.

Eicher, T. S., A. Lenkoski, and A. Raftery, 2010, Bayesian model averaging and endogeneity under model uncertainty: An application to development determinants, Econometric Reviews (forthcoming).

Eicher, T. S., A. Lenkoski, and A. E. Raftery, 2009, Bayesian Model Averaging andEndogeneity Under Model Uncertainty: An Application to Development Determinants,Working Paper, University of Washington.

Engerman, S., and K. Sokoloff, 2002, Factor Endowments, Inequality and Paths of Development Among New World Economies, NBER Working Papers, 9259.

Fernandez, C., E. Ley, and M. Steel, 2001, Model Uncertainty in Cross-Country Growth Regressions, Journal of Applied Econometrics 16, 563–576.

Frankema, Ewout, 2009, The Colonial Roots of Land Inequality: Geography, Factor Endowments, or Institutions?, Economic History Review 63, 418–451.

Gallego, F. A., 2010, Historical Origins of Schooling: The Role of Democracy and Political Decentralization, The Review of Economics and Statistics 92, 228–243.

Galor, O., and O. Moav, 2006, Das Human Kapital: A Theory of the Demise of the Class Structure, Review of Economic Studies 73, 85–117.

Galor, O., and D. Vollrath, 2009, Inequality in Land Ownership, the Emergence of Human Capital Promoting Institutions, and the Great Divergence, Review of Economic Studies 76, 143–179.

Galor, O., and D. Weil, 2000, Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond, American Economic Review 90, 806–828.

Glaeser, E. L., R. La Porta, F. Lopez-De-Silanes, and A. Shleifer, 2004, Do Institutions Cause Growth, Journal of Economic Growth 9, 271–303.

Goldin, C., and L.F. Katz, 1997, Why the United States Led in Education: Lessons from Secondary School Expansion, 1910 to 1940, NBER Working Papers, 6144.

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.

Hansen, B., 1996, Inference when a nuisance parameter is not identified under the null hypothesis, Econometrica 64, 413–430.

Hansen, B., 2000, Sample splitting and threshold estimation, Econometrica 68, 575–603.

Hansen, G., and E. C. Prescott, 2001, Malthus to Solow, American Economic Review 92,1205–17.

Jeffreys, H., 1961, Theory of Probability (Oxford University Press).

Kass, R. E., and A. E. Raftery, 1995, Bayes factors, Journal of the American Statistical Association 90, 773-795.

Koren, M., and S Tenreyro, 2007, Volatility and development, Quarterly Journal of Economics 122, 243–287.

Kourtellos, A., A. Stengos, and C. M. Tan, 2011, Structural threshold regression, Working paper, University of Cyprus, Clark University, and University of Guelph.

Kourtellos, A., C. M. Tan, and X. Zhang, 2007, Is the relationship between aid and economic growth nonlinear?, Journal of Macroeconomics 29, 515–540.

La Porta, R., F. Lopez-de Silanes, A. Shleifer, and R. W. Vishny, 1999, The Quality of Government, Journal of Law, Economics and Organization 15, 222–279.

Leamer, E. E., 1978, Specification Searches (Wiley, New York).

Maddison, A., 2009, Statistics on World Population, GDP, and Per Capita GDP, 1-2006 AD.

Madigan, D., and A.E Raftery, 1994, Model selection and accounting for model uncertainty in graphical models using Occam's Window, Journal of the American Statistical Association, 89, 1335-1346.

Malik, A., and J.R.W. Temple, 2009, The geography of output volatility, Journal of Development Economics 90, 163–178.

Mamuneas, T., A. Savvides, and T. Stengos, 2006, Economic development and the return to human capital: A smooth coefficient semiparametric approach, Journal of Applied Econometrics 21, 111–132.

Masanjala, W., and C. Papageorgiou, 2008, Rough and Lonely Road to Prosperity: A Reexamination of the Sources of Growth in Africa Using Bayesian Model Averaging, Journal of Applied Econometrics 23, 671–682.

Miller, M. K., 2011, Democracy by Example? Economic Growth, Policy Diffusion, and Regime Change., Princeton, mimeo.

Mitchell, B.R, 1998, International Historical Statistics: The Americas, Africa, Asia, Oceania, and Europe 1750-1993 (MacMillan Press).

Mobarak, A. M., 2005, Democracy, volatility, and economic development, The Review of Economics and Statistics 87, 348–361.

Papageorgiou, C., 2002, Trade as a threshold variable for multiple regimes, Economics Letters 77, 85–91.

Parente, S., and E. Prescott, 2000, Barriers to Riches (MIT Press).

Perron, P., 1989, The Great Crash, the Oil Price Shock and the Unit Root Hypothesis, Econometrica, 57, 1361-1401.

Perron, P. and D. Kim, 2009, Unit root tests allowing for a break in the trend function at an unknown time under both the null and alternative hypotheses, Journal of Econometrics, 148, 1-13.

Perron, P and Z. Qu, 2007, Estimating and Testing Structural Changes in Multivariate Regressions, Econometrica, 75(2), 459-502.

Perron, P. and T. Yabu, 2009, Testing for Shifts in Trend With an Integrated or Stationary Noise Component, Journal of Business and Economic Statistics, 27(3), 369-396.

Qu, Z., 2007, Searching for co-integration in a dynamic system, Econometrics Journal, 10, 580-604.

Raftery, A., D. Madigan, and J. Hoeting, 1997, Bayesian model averaging for linear regression models, Journal of the American Statistical Association 92, 179-191.

Rajan, R. G., and R. Ramcharan, 2010, Land and Credit: A Study of the Political Economy of Banking in the United States in the Early 20th Century, Journal of Finance (forthcoming).

Ramankutty, N., and J. A. Foley, 1999, Estimating historical changes in global land cover: Croplands from 1700 to 1992, Global Biogeochemical Cycles 13, 997–1027.

Ramey, G., and V. Ramey, 1995, Cross-country evidence on the link between volatility and growth, American Economic Review 85, 1138–1151.

Sachs, J., J. L. Gallup, and A. Mellinger, 1999, Geography and Economic Development, International Regional Science Review 22, 172–232.

Sala-i Martin, X., G. Doppelhofer, and R. Miller, 2004, Determinants of Long-term Growth: a Bayesian Averaging of Classical Estimates (BACE) Approach, American Economic Review 94, 813–835.

Sims, C.A., 1980, Macroeconomics and reality, Econometrica , 48, 1-48.

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.

Van Zanden, J.L, J. Baten, P. Foldvari, and B. Van Leeuwen, 2011, The Changing Shape of Global Inequality: Exploring a new dataset, Center for Global Economic History (CGEH), No. 1, CGEH working paper series.

Vogelsang, T. J. and P. Perron, 1998, Additional tests for a unit root allowing the possibility of breaks in the trend function, International Economic Review, 39, 1073-1100.

Volinsky, C., D. Madigan, A. Raftery, and R. Kronmal, 1997, Bayesian Model Averaging in Proportional Hazards Model: Predicting the Risk of a Stroke, Applied Statistics 46, 443–448.

Zivot, E. and D. W. K. Andrews, 1992, Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit Root Hypothesis, Journal of Business and Economic Statistics, 10 (3), 251-270.